
SIGILLVM · VNIVERSITATIS · CALIFORNIENSIS
LVX
FIAT
MDCCCLXVIII
EX LIBRIS

ELEMENTARY THEORY
OF
EQUATIONS

BY

LEONARD EUGENE DICKSON, PH.D.

CORRESPONDANT DE L'INSTITUT DE FRANCE
PROFESSOR OF MATHEMATICS IN THE
UNIVERSITY OF CHICAGO

NEW YORK
JOHN WILEY & SONS, INC.
LONDON: CHAPMAN & HALL, LIMITED

Stanhope Press
F. H. GILSON COMPANY
BOSTON, U. S. A.

4-35

PREFACE

The longer an engineer has been separated from his alma mater, the fewer mathematical formulas he uses and the more he relies upon tables and, when the latter fail, upon graphical methods. Although graphical methods have the advantage of being ocular, they frequently suffer from the fact that only what is seen is sensed. But this defect is due to the kind of graphics used. With the aid of the scientific art of graphing presented in Chapter I, one may not merely make better graphs in less time but actually draw correct negative conclusions from a graph so made, and therefore sense more than one sees. For instance, one may be sure that a given cubic equation has only the one real root seen in the graph, if the bend points lie on the same side of the x-axis.

Emphasis is here placed upon Newton's method of solving numerical equations, both from the graphical and the numerical standpoint. One of several advantages (well recognized in Europe) of Newton's method over Horner's is that it applies as well to non-algebraic as to algebraic equations.

In this elementary book, the author has of course omitted the difficult Galois theory of algebraic equations (certain texts on which are very erroneous) and has merely illustrated the subject of invariants by a few examples.

It is surprising that the theorems of Descartes, Budan, and Sturm, on the real roots of an equation, are often stated inaccurately. Nor are the texts in English on this subject more fortunate on the score of correct proofs; for these reasons, care has been taken in selecting the books to which the reader is referred in the present text.

The material is here so arranged that, before an important general theorem is stated, the reader has had concrete illustrations and often also special cases. The exercises are so placed that a reasonably elegant and brief solution may be expected, without resort to tedious multiplications and similar manual labor. Very few of the five hundred exercises are of the same nature.

Complex numbers are introduced in a logical and satisfying manner. The treatment of roots of unity is concrete, in contrast to the usual abstract method.

Attention is paid to scientific computation, both as to control of the limit of error and as to securing maximum accuracy with minimum labor.

An easy introduction to determinants and their application to the solution of systems of linear equations is afforded by Chapter XI, which is independent of the earlier chapters.

Here and there are given brief, but clear, outlooks upon various topics of decided intrinsic and historical interest, — thus putting real meat upon the dry bones of the subject.

To provide for a very brief course, certain sections, aggregating over fifty pages, are marked by a dagger for omission. However, in compensation for the somewhat more advanced character of these sections, they are treated in greater detail.

In addition to the large number of illustrative problems solved in the text, there are five hundred very carefully selected and graded exercises, distributed into seventy sets. As only sixty of these exercises (falling into seventeen sets) are marked with a dagger, there remains an ample number of exercises for the briefer course.

The author is greatly indebted to his colleagues Professors A. C. Lunn and E. J. Wilczynski for most valuable suggestions made after reading the initial manuscript of the book. Useful advice was given by Professor G. A. Miller, who read part of the galley proofs. A most thorough reading of both the galley and page proofs was very generously made by Dr. A. J. Kempner, whose scientific comments and very practical suggestions have led to a marked improvement of the book. Moreover, the galleys were read critically by Professor D. R. Curtiss, who gave the author the benefit not merely of his wide knowledge of the subject but also of his keen critical ability. The author sends forth the book thus emended with less fear of future critics, and with the hope that it will prove as stimulating and useful as these five friends have been generous of their aid.

CHICAGO, *February*, 1914.

CONTENTS

THEORY OF EQUATIONS

CHAPTER I

The Graph of an Equation

1. For purposes of review, certain terms already familiar to the reader are defined here. Through a point O, called the *origin*, draw a horizontal straight line OX and a vertical straight line OY. These lines are called the *axes of coördinates;* in particular, OX is called the *x-axis*. Choose a convenient unit of length. Consider any point P in the plane and let Q be the point of intersection of the x-axis with the vertical line through P. By the *abscissa* x of the point P is understood the number of units of length in OQ in case Q lies to the right of the origin O; but, in case Q lies to the left of O, x is the negative of the number of units of length in OQ. Similarly, the *ordinate* y of the point P is the length of PQ when P is above the x-axis, but is the negative of the length of PQ when P is below the x-axis. For the point P in Fig. 1, $x = +4$, $y = -11$. The real numbers x and y which a point determines in this manner are called its *coördinates*. Conversely, any pair of real numbers determines a point.

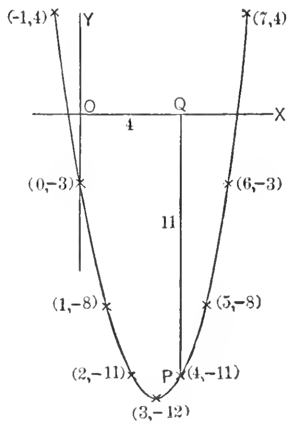


Fig. 1

Figure 1 shows the points which represent various pairs of values of x and y, satisfying the equation

$$(1) \qquad y = x^2 - 6x - 3.$$

For example, the point P represents the pair of values $x = 4$, $y = -11$, and is designated $(4, -11)$. Since the value of x may be assigned at pleasure and a corresponding value of y is determined by equation (1), there is an infinitude of points representing pairs of values

satisfying the equation. These points constitute a curve called the *graph* of the equation.

In Fig. 1, the curve intersects the x-axis in two points; the abscissa of one point of intersection is between 6 and 7, that of the other point is between -1 and 0. The x-axis is the graph of the equation $y = 0$. Thus the abscissas of the intersections of the graph of equation (1) and the graph of $y = 0$ are the real roots of the quadratic equation

$$(1') \qquad x^2 - 6x - 3 = 0.$$

Hence to find graphically the real roots of the last equation, we equate the left member to y and use the graph of the resulting equation (1). For other methods, see §§ 16–18.

EXERCISES

1. Find graphically the real roots of $x^2 - 6x + 7 = 0$.
2. Discuss graphically the reality of the roots of $x^2 - 6x + 12 = 0$.
3. Obtain the graph used in Ex. 1 by shifting the graph in Fig. 1 ten units upwards, leaving the axes OX and OY unchanged. How may we obtain similarly that used in Ex. 2?
4. Locate graphically the real roots of $x^3 + 4x^2 - 7 = 0$.

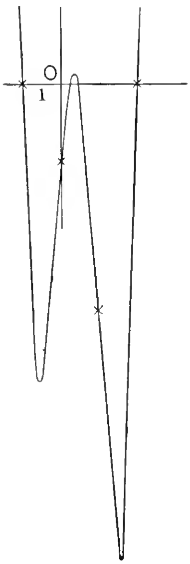


Fig. 2

2. Caution in Plotting. If the example set were

$$(2) \qquad y = 8x^4 - 14x^3 - 9x^2 + 11x - 2,$$

one might use successive integral values of x, obtain the points $(-2, 180)$, $(-1, 0)$, $(0, -2)$, $(1, -6)$, $(2, 0)$, $(3, 220)$, all but the first and last of which are shown (by crosses) in Fig. 2, and be tempted to conclude that the graph is a U-shaped curve approximately like that in Fig. 1 and that there are just two real roots, -1 and 2, of

$$(2') \qquad 8x^4 - 14x^3 - 9x^2 + 11x - 2 = 0.$$

But both of these conclusions would be false. In fact, the graph is a W-shaped curve (Fig. 2) and the additional real roots are $\frac{1}{4}$ and $\frac{1}{2}$.

This example shows that it is often necessary to employ also values of x which are not integers. The purpose of the example was, however, not to point out this obvious fact, but rather to emphasize the chance of serious error in sketching a curve

through a number of points, however numerous. The true curve between two points below the x-axis may not cross the x-axis, or may have a peak actually crossing the x-axis twice, or may be an M-shaped curve crossing it four times, etc.

For example, the graph (Fig. 3) of

$$y = x^3 + 4x^2 - 11 \tag{3}$$

crosses the x-axis only once. But this fact can not be concluded from a graph located by a number of points, however numerous, whose abscissas are chosen at random.

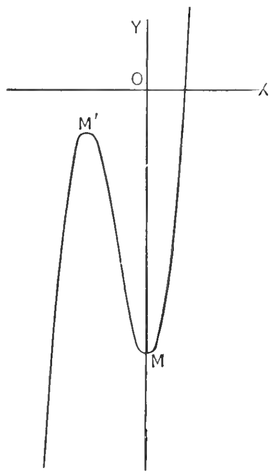


Fig. 3

We shall find that correct conclusions regarding the number of real roots can be deduced from a graph whose bend points (§ 3) have been located.

We shall be concerned with equations of the form

$$a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = 0 \qquad (a_0 \neq 0),$$

in which $a_0, a_1, \ldots a_n$ are real constants. The left member is called a *polynomial* in x of *degree* n, or also a *rational integral function* of x, and will frequently be denoted for brevity by the symbol $f(x)$ and less often by f.

3. Bend Points. A point (like M or M' in Fig. 3) is called a *bend point* of the graph of $y = f(x)$ if the tangent to the graph at that point is horizontal and if all of the adjacent points of the graph lie below the tangent or all above the tangent. The first, but not the second, condition is satisfied by the point O of the graph of $y = x^3$ given in Fig. 4 (see § 6). In the language of the calculus, $f(x)$ has a (relative) maximum or minimum value at the abscissa of a bend point on the graph of $y = f(x)$.

Let $P = (x, y)$ and $Q = (x + h, Y)$ be two points on the graph, sketched in Fig. 5, of $y = f(x)$. By the *slope* of a straight line is meant

the tangent of the angle between the line and the x-axis measured counter-clockwise from the latter. In Fig. 5, the slope of the straight line PQ is

$$\frac{Y-y}{h}=\frac{f(x+h)-f(x)}{h}. \tag{4}$$

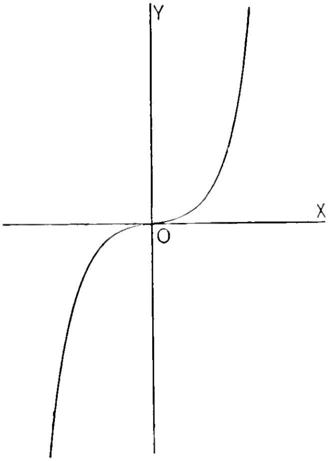


Fig. 4

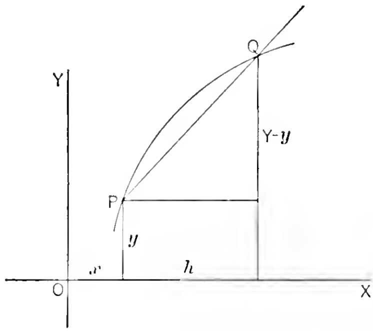


Fig. 5

For equation (3), $f(x) = x^3 + 4x^2 - 11$. Hence

$$\begin{aligned} f(x+h) &= (x+h)^3 + 4(x+h)^2 - 11 \\ &= x^3 + 4x^2 - 11 + (3x^2 + 8x)h + (3x+4)h^2 + h^3. \end{aligned}$$

The slope (4) of the secant PQ is here

$$3x^2 + 8x + (3x+4)h + h^2.$$

Now let the point Q move along the graph towards P. Then h approaches the value zero and the secant PQ approaches the tangent at P. The slope of the tangent at P is therefore the corresponding limit $3x^2 + 8x$ of the preceding expression.

In particular, if P is a bend point the slope of the tangent at P is zero and hence $x = 0$ or $x = -\frac{8}{3}$. Equation (3) gives the corresponding values of y. The resulting points

$$M = (0, -11), \quad M' = (-\tfrac{8}{3}, -\tfrac{41}{27})$$

are easily shown to be bend points. Indeed, for $x > 0$ and for x between -4 and 0, $x^2(x + 4)$ is positive, and hence $f(x) > -11$ for such values of x, so that the function (3) has a relative minimum at $x = 0$. Similarly, there is a relative maximum at $x = -\frac{8}{3}$. We may also employ the general method of § 8 to show that M and M' are bend points. Since these bend points are both below the x-axis, we are now certain that the graph crosses the x-axis only once.

The use of the bend points insures greater accuracy to the graph than the use of dozens of points whose abscissas are taken at random.

4. Derivatives. We shall now find the slope of the tangent to the graph of $y = f(x)$, where $f(x)$ is any polynomial

$$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n. \tag{5}$$

We need the expansion of $f(x + h)$ in powers of x. By the binomial theorem,

$$a_0(x + h)^n = a_0x^n + na_0x^{n-1}h + \frac{n(n-1)}{2}a_0x^{n-2}h^2 + \cdots,$$

$$a_1(x + h)^{n-1} = a_1x^{n-1} + (n-1)a_1x^{n-2}h + \frac{(n-1)(n-2)}{2}a_1x^{n-3}h^2 + \cdots,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$a_{n-2}(x + h)^2 = a_{n-2}x^2 + 2\,a_{n-2}xh + a_{n-2}h^2,$$

$$a_{n-1}(x + h) = a_{n-1}x + a_{n-1}h,$$

$$a_n = a_n.$$

The sum of the left members is evidently $f(x + h)$. On the right, the sum of the first terms (*i.e.*, those free of h) is $f(x)$. The sum of the coefficients of h is denoted by $f'(x)$, the sum of the coefficients of $\frac{1}{2}h^2$ is denoted by $f''(x)$, $\cdots$, the sum of the coefficients of

$$\frac{h^k}{1 \cdot 2 \cdots k}$$

is denoted by $f^{(k)}(x)$. Thus

$$f'(x) = na_0x^{n-1} + (n-1)a_1x^{n-2} + \cdots + 2\,a_{n-2}x + a_{n-1}, \tag{6}$$

$$f''(x) = n(n-1)\,a_0x^{n-2} + (n-1)(n-2)a_1x^{n-3} + \cdots + 2\,a_{n-2}, \tag{7}$$

etc. Hence we have

$$f(x + h) = f(x) + f'(x)\,h + f''(x)\frac{h^2}{1 \cdot 2} + f'''(x)\frac{h^3}{1 \cdot 2 \cdot 3} + \cdots + f^{(n)}(x)\frac{h^n}{1 \cdot 2 \cdots n}. \tag{8}$$

This formula (8) is known as *Taylor's theorem* for the present case of a polynomial $f(x)$ of degree n. We call $f'(x)$ the *(first) derivative* of $f(x)$, and $f''(x)$ the *second derivative* of $f(x)$, etc. Concerning the fact that $f''(x)$ is the first derivative of $f'(x)$ and that, in general, the kth derivative $f^{(k)}(x)$ of $f(x)$ equals the first derivative of $f^{(k-1)}(x)$, see Exs. 6–9 of the next set.

In view of (8), the limit of (4) as h approaches zero is $f'(x)$. Hence *$f'(x)$ is the slope of the tangent to the graph of $y = f(x)$ at the point (x, y).*

In (5) and (6), let every a be zero except a_0. Thus the derivative of a_0x^n is na_0x^{n-1}, and hence is obtained by multiplying the given term by its exponent n and then diminishing its exponent by unity. For example, the derivative of $2x^3$ is $6x^2$.

Moreover, the derivative of $f(x)$ equals the sum of the derivatives of its separate terms. Thus the derivative of $x^3 + 4x^2 - 11$ is $3x^2 + 8x$, as found also in § 3.

5. Computation of Polynomials. The labor of computing the value of a polynomial $f(x)$ for a given value of x may be much shortened by a simple device. To find the value of

$$x^3 + 3x^2 - 2x - 5$$

for $x = 2$, we note that $x^3 = x \cdot x^2 = 2x^2$, so that the sum of the first two terms is $5x^2$. This latter equals $5 \cdot 2x$ or $10x$, adding this to the next term $-2x$, we get $8x$ or 16. The final result is therefore **11**.

Write the coefficients in a line. Then the work is:

$$\begin{array}{rrrr|l} 1 & 3 & -2 & -5 & \underline{2} \\ & 2 & 10 & 16 & \\ \hline 1 & 5 & 8 & 11. & \end{array}$$

In case not all the intermediate powers of x occur among the terms of $f(x)$, the missing powers are considered as having the coefficients zero. Thus the value -61 of $2x^5 - x^3 + 2x - 1$ for $x = -2$ is found as follows:

$$\begin{array}{rrrrrr|l} 2 & 0 & -1 & 0 & 2 & -1 & \underline{-2} \\ & -4 & 8 & -14 & 28 & -60 & \\ \hline 2 & -4 & 7 & -14 & 30 & -61. & \end{array}$$

For another manner of presenting this method see Ch. X, § 4.

EXERCISES

1. The slope of the tangent to $y = 8x^3 - 22x^2 + 13x - 2$ at (x, y) is $24x^2 - 44x + 13$. The bend points are $(0.37, 0.203)$, $(1.46, -5.03)$, approximately. Draw the graph.

2. The bend points of $y = x^3 - 2x - 5$ are $(.82, -6.09)$, $(-.82, -3.91)$, approximately. Draw the graph and locate the real roots.

3. Find the bend points of $y = x^3 + 6x^2 + 8x + 8$. Locate the real roots.

4. Locate the real roots of $f(x) = x^4 + x^3 - x - 2 = 0$. The abscissas of the bend points are the roots of $f'(x) = 4x^3 + 3x^2 - 1 = 0$. The bend points of $y = f'(x)$ are $(0, -1)$ and $(-\frac{1}{2}, -\frac{3}{4})$, so that $f'(x) = 0$ has a single real root (it is just less than $\frac{1}{2}$). The single bend point of $y = f(x)$ is $(\frac{1}{2}, -\frac{37}{16})$, approximately.

5. Locate the real roots of $x^6 - 7x^4 - 3x^2 + 7 = 0$.

6. $f''(x)$, given by (7), is the first derivative of $f'(x)$.

7. If $f(x) = f_1(x) + f_2(x)$, the kth derivative of f equals the sum of the kth derivatives of f_1 and f_2. Use (8).

8. $f^{(k)}(x)$ equals the first derivative of $f^{(k-1)}(x)$. Hint: prove this for $f = ax^m$; then prove that it is true for $f = f_1 + f_2$ if true for f_1 and f_2.

9. Find the third derivative of $x^6 + 5x^4$ by forming successive first derivatives; also that of $2x^5 - 7x^3 + x$.

10. The derivative of gk is $g'k + gk'$. Hint: multiply the members of $g(x+h) = g(x) + g'(x)h + \cdots$ and $k(x+h) = k(x) + k'(x)h + \cdots$ and use (8) for $f = gk$.

6. Horizontal Tangents. If (x, y) is a bend point of the graph of $y = f(x)$, then, by definition, the slope of the tangent at (x, y) is zero. Hence (§ 4), the abscissa x is a root of $f'(x) = 0$. In Exs. 1–5 of the preceding set, it was true that, conversely, any real root of $f'(x) = 0$ is the abscissa of a bend point. However, this is not always the case. We shall now consider in detail an example illustrating this fact. The example is the one merely mentioned in § 3 to indicate the need of the second requirement made in our definition of a bend point.

The graph (Fig. 4) of $y = x^3$ has no bend point since x^3 increases when x increases. Nevertheless, the derivative $3x^2$ of x^3 is zero for the real value $x = 0$. The tangent to the curve at $(0, 0)$ is the horizontal line $y = 0$. It may be thought of as the limiting position of a secant through O which meets the curve in two further points, seen to be equidistant from O. When one, and hence also the other, of the latter points approaches O, the secant approaches the position of tangency. In this sense the tangent at O is said to meet the curve in three coincident points, their abscissas being the three coinciding roots of $x^3 = 0$. In the

usual technical language which we shall employ henceforth, $x^3 = 0$ has the *triple root* $x = 0$. The subject of bend points, to which we recur in § 8, has thus led us to a digression on the important subject of double roots, triple roots, etc.

7. Multiple Roots. In (8) replace x by α and h by $x - \alpha$. Then

$$(9) \qquad f(x) = f(\alpha) + f'(\alpha)(x-\alpha) + f''(\alpha)\frac{(x-\alpha)^2}{1\cdot 2} + f'''(\alpha)\frac{(x-\alpha)^3}{1\cdot 2\cdot 3} + \cdots.$$

Thus the constant remainder obtained by dividing any polynomial $f(x)$ by $x - \alpha$ is $f(\alpha)$, a fact known as the Remainder Theorem. In particular, if $f(\alpha) = 0$, $f(x)$ has the factor $x - \alpha$. This proves the Factor Theorem: If α is a root of $f(x) = 0$, then $x - \alpha$ is a factor of $f(x)$.

The converse is true: If $x - \alpha$ is a factor of $f(x)$, then α is a root of $f(x) = 0$. In case $f(x)$ has the factor $(x - \alpha)^2$, but not the factor $(x - \alpha)^3$, α is called a *double* root of $f(x) = 0$. In general, if $f(x)$ has the factor $(x - \alpha)^m$, but not the factor $(x - \alpha)^{m+1}$, α is called a *multiple* root of *multiplicity* m of $f(x) = 0$, or an m-fold root. Thus, 4 is a simple root, 3 a double root and -2 a triple root of

$$7\,(x-4)(x-3)^2(x+2)^3 = 0.$$

This algebraic definition of a multiple root is in fact equivalent to the geometrical definition, given for a special case, in § 6.

The second member of (9) is divisible by $(x - \alpha)^2$ if and only if $f(\alpha) = 0$, $f'(\alpha) = 0$, and is divisible by $(x - \alpha)^3$ if and only if also $f''(\alpha) = 0$, etc. Hence *α is a double root of $f(x) = 0$ if and only if $f(\alpha) = 0$, $f'(\alpha) = 0$, $f''(\alpha) \neq 0$; α is a root of multiplicity m if and only if*

$$(10) \qquad f(\alpha) = 0,\ f'(\alpha) = 0,\ f''(\alpha) = 0,\ \cdots,\ f^{(m-1)}(\alpha) = 0,\ f^{(m)}(\alpha) \neq 0.$$

For example, zero is a triple root of $x^4 + 2\,x^3 = 0$ since the first and second derivatives are zero for $x = 0$, while the third derivative $24\,x + 12$ is not.

If $f(x)$ and $f'(x)$ have the common factor $(x - \alpha)^{m-1}$, but not $(x - \alpha)^m$, where $m \geqq 2$, then α is a root of $f(x) = 0$ of multiplicity m. For, α is a root of multiplicity at least $m - 1$ of both $f(x) = 0$ and $f'(x) = 0$, so that the equalities in (10) hold; also $f^{(m)}(\alpha) \neq 0$ holds, since otherwise α would be a root of both $f(x) = 0$ and $f'(x) = 0$ of multiplicity m or greater, and $(x - \alpha)^m$ would be a common factor. Hence *if $f(x)$ and $f'(x)$ have a greatest common divisor $g(x)$ involving x, a root of $g(x) = 0$ of multiplicity*

$m-1$ is a root of $f(x)=0$ of multiplicity m, and conversely any root of $f(x)=0$ of multiplicity m is a root of $g(x)=0$ of multiplicity $m-1$. The last fact follows from relations (10), which imply that α is a root of $f'(x)=0$ of multiplicity $m-1$, and hence that $f(x)$ and $f'(x)$ have the common factor $(x-\alpha)^{m-1}$, but not $(x-\alpha)^m$.

In view of this theorem, the problem of finding all the multiple roots of $f(x)=0$ and the multiplicity of each multiple root is reduced to the problem of finding the roots of $g(x)=0$ and the multiplicity of each.

For example, let $f(x) = x^3 - 2x^2 - 4x + 8$. Then

$$f'(x) = 3x^2 - 4x - 4, \quad 9f(x) = f'(x)\,(3x-2) - 32\,(x-2).$$

Since $x-2$ is a factor of $f'(x)$ it may be taken to be the greatest common divisor of $f(x)$ and $f'(x)$, as the choice of the constant factor c in $c(x-2)$ is here immaterial. Hence 2 is a double root of $f(x)=0$, while the remaining root -2 is a simple root.

EXERCISES

1. $x^3 - 7x^2 + 15x - 9 = 0$ has a double root.
2. $x^4 - 8x^2 + 16 = 0$ has two double roots.
3. $x^4 - 6x^2 - 8x - 3 = 0$ has a triple root.
4. Test $x^4 - 8x^3 + 22x^2 - 24x + 9 = 0$ for multiple roots.
5. Test $x^3 - 6x^2 + 11x - 6 = 0$ for multiple roots.

8. Inflexion and Bend Points. The equation of the tangent to the graph of $y=f(x)$ at the point (α, β) on it is

$$y = f'(\alpha)\,(x-\alpha) + \beta \qquad [\beta = f(\alpha)].$$

For the abscissas of its intersections with the graph of $y=f(x)$, we have, from (9),

$$f''(\alpha)\frac{(x-\alpha)^2}{1\cdot 2} + f'''(\alpha)\frac{(x-\alpha)^3}{1\cdot 2\cdot 3} + \cdots = 0.$$

If α is a root of multiplicity m of this equation, the point (α, β) is counted as m coincident points of intersection of the tangent and the curve (just as in the example in § 6). This will be the case if and only if *

$$(11) \qquad f''(\alpha)=0,\ f'''(\alpha)=0,\ \ldots,\ f^{(m-1)}(\alpha)=0,\ f^{(m)}(\alpha)\neq 0.$$

For example, if $f(x) = x^4$ and $\alpha = 0$, then $m = 4$. The graph of $y = x^4$ is a U-shaped curve, whose intersection with the tangent (x-axis) at (0, 0) is counted as four coincident points of intersection.

* If $m = 2$, only the last relation of the set is retained: $f''(\alpha) \neq 0$.

If m is even, the points of the curve in the vicinity of the point of tangency (α, β) are all on the same side of the tangent and the point (α, β) is, by the definition in § 3, a bend point. But if m is odd $(m > 1)$, the curve crosses the tangent at the point of tangency (α, β) and this point is called an *inflexion point*, and the tangent an *inflexion tangent*. To simplify the proof, take (α, β) as the new origin of coördinates and the tangent as the new x-axis. Then the new equation of the curve is

$$y = cx^m + dx^{m+1} + \cdots \qquad (c \neq 0,\ m \geqq 2).$$

For x sufficiently small numerically, y has the same sign as cx^m (§ 11). Thus if m is even, the points of the curve in the vicinity of the origin are all on the same side of the x-axis. But if m is odd, the points with small positive abscissas lie on one side of the x-axis and those with numerically small negative abscissas lie on the opposite side.

For example, $(0, 0)$ is a bend point of the graph of $y = x^4$. But $(0, 0)$ is an inflexion point of the graph (Fig. 4) of $y = x^3$, and the inflexion tangent $y = 0$ crosses the curve at $(0, 0)$. Here $f''(0) = 0$, $f'''(0) = 6$, so that $m = 3$, in accord with the evident fact that $x^3 = 0$ has the root zero of multiplicity 3.

We have, therefore, in the evenness or oddness of m in (11) a practical test to decide which roots α of $f'(x) = 0$ are abscissas of bend points and which are abscissas of inflexion points with horizontal inflexion tangents.

EXERCISES

1. If $f(x) = 3x^5 + 5x^3 + 4$, the only real root of $f'(x) = 0$ is $x = 0$. Show that $(0, 4)$ is an inflexion point, and thus that there is no bend point and hence that $f(x) = 0$ has a single real root.

2. $x^3 - 3x^2 + 3x + c = 0$ has an inflexion point, but no bend point.

3. $x^5 - 10x^3 - 20x^2 - 15x + c = 0$ has two bend points and no horizontal inflexion tangents.

4. $3x^5 - 40x^3 + 240x + c = 0$ has no bend point, but has two horizontal inflexion tangents.

5. Any function $x^3 - 3\alpha x^2 + \cdots$ of the third degree can be written in the form $f(x) = (x - \alpha)^3 + ax + b$. The straight line having the equation $y = ax + b$ meets the graph of $y = f(x)$ in three coincident points with the abscissa α and hence is an inflexion tangent. If we take new axes of coördinates parallel to the old and intersecting at the new origin $(\alpha, 0)$, *i.e.*, if we make the transformation $x = X + \alpha$, $y = Y$, of coördinates, we see that the equation $f(x) = 0$ becomes a reduced cubic equation $X^3 + pX + q = 0$ (cf. Ch. III).

6. Find the inflexion tangent to $y = x^3 + 6x^2 - 3x + 1$ and transform $x^3 + 6x^2 - 3x + 1 = 0$ into a reduced cubic equation.

9. Real Roots of a Cubic Equation. It suffices to consider

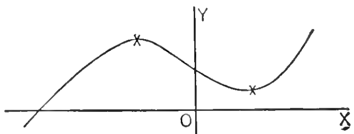
$$f(x) = x^3 - 3\,lx + q \qquad (l \neq 0),$$

in view of Ex. 5 above. Then $f' = 3\,(x^2 - l)$, $f'' = 6\,x$. If $l < 0$, there is no bend point and the cubic equation $f(x) = 0$ has a single real root.

If $l > 0$, there are two bend points

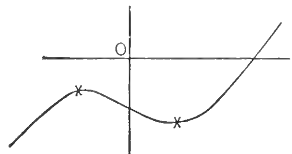
$$(\sqrt{l},\, q - 2\,l\sqrt{l}),\ (-\sqrt{l},\, q + 2\,l\sqrt{l})$$

and the graph of $y = f(x)$ is evidently of one of the three types:



$q \geqq 2\,l\sqrt{l}$

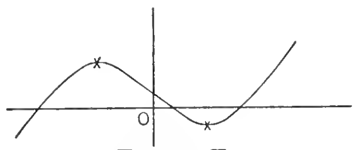
Fig. 6



$q \leqq -2\,l\sqrt{l}$

Fig. 7

If the equality sign holds in the first or second case, one of the bend points is on the x-axis and the cubic equation has a double root; the condition is that $q^2 - 4\,l^3 = 0$. The third case is fully specified by the condition $q^2 < 4\,l^3$, which implies that $l > 0$. Hence *$x^3 - 3\,lx + q = 0$ has three distinct real roots if and only if $q^2 < 4\,l^3$, a single real root if and only if $q^2 > 4\,l^3$; and a double root (necessarily real) if and only if $q^2 = 4\,l^3$.*



$-2\,l\sqrt{l} < q < 2\,l\sqrt{l}$

Fig. 8

EXERCISES.

Apply the criterion to find the number of real roots of:

1. $x^3 + 2\,x - 4 = 0$.
2. $x^3 - 7\,x + 7 = 0$.
3. $x^3 - 2\,x - 1 = 0$.
4. $x^3 - 3\,x + 2 = 0$.
5. $x^3 + 6\,x^2 - 3\,x + 1 = 0$.
6. The inflexion point of $y = x^3 - 3\,lx + q$ is $(0, q)$.

10.† Trinomial Equations.

For m and n positive odd integers, $m > n$, let

$$f(x) = x^m + px^n + q \qquad (p \neq 0).$$

Here $x = 0$ is a root of $f'(x) = 0$ only when $n > 1$ and then the tangent at $(0, q)$ is the horizontal inflexion tangent $y = q$, as shown by (11) with m replaced by n, or directly from the fact that zero is a root of odd multiplicity n of $x^m + px^n = 0$. Hence in no case is zero the abscissa of a bend point.

If $p > 0$, f' has no real root except $x = 0$. Thus there is no bend point and hence a single real root of $f(x) = 0$.

If $p < 0$, there are just two bend points, their abscissas being b and $-b$, where b is the single positive real root of $b^{m-n} = -np/m$. The bend points are on the same side or opposite sides of the x-axis according as

$$f(b) = q + pb^n\left(1 - \frac{n}{m}\right), \qquad f(-b) = q - pb^n\left(1 - \frac{n}{m}\right)$$

are of like signs or opposite signs. The number of real roots is 1 or 3 in the respective cases. Hence there are three distinct real roots if and only if the positive number

$$-pb^n\left(1 - \frac{n}{m}\right)$$

exceeds both q and $-q$, *i.e.*, if

$$-p\frac{n}{m}b^n > \frac{\pm nq}{m - n}.$$

The first member equals b^m, so that its $(m - n)$th power is the mth power of $b^{m-n} = -np/m$. Hence the conditions are equivalent to

$$0 > \left(\frac{np}{m}\right)^m + \left(\frac{nq}{m - n}\right)^{m-n}.$$

EXERCISES†

1.† $x^3 + px + q = 0$ has three distinct real roots if and only if

$$0 > \left(\frac{p}{3}\right)^3 + \left(\frac{q}{2}\right)^2.$$

2.† If p and q are positive, $x^{2m} - px^{2n} + q = 0$ has four distinct real roots, two pairs of equal roots, or no real root, according as

$$\left(\frac{np}{m}\right)^m - \left(\frac{nq}{m - n}\right)^{m-n} > 0, \ = 0, \text{ or } < 0.$$

11. Continuity of a Polynomial. Hitherto we have located certain points of the graph of $y = f(x)$, where $f(x)$ is a polynomial in x with real coefficients, and taken the liberty to join them by a continuous curve.

The polynomial $f(x)$ in the real variable x shall be called *continuous at* $x = a$, where a is a real constant, if the difference

$$D = f(a + h) - f(a)$$

is numerically less than any assigned positive number p for all real values of h sufficiently small numerically.

We shall prove that any polynomial $f(x)$ with real coefficients is continuous at $x = a$, where a is any real constant.

The proof rests upon Taylor's formula (8), which gives

$$D = f'(a)h + \frac{f''(a)}{1 \cdot 2} h^2 + \cdots + \frac{f^{(n)}(a)}{1 \cdot 2 \cdots n} h^n.$$

Denote by g the greatest numerical value of the coefficients of h, $h^2, \ldots, h^n$. For h numerically less than k, where $k < 1$, we see that D is numerically less than

$$g(k + k^2 + \cdots + k^n) < g\frac{k}{1 - k} < p, \quad \text{if } k < \frac{p}{p + g}.$$

The same proof shows that, if $a_1, \ldots, a_n$ are real, $a_1h + \cdots + a_nh^n$ is numerically less than an assigned positive number p for all real values of h sufficiently small numerically.

12. Theorem. *If the coefficients of the polynomial $f(x)$ are real and if a and b are real numbers such that $f(a)$ and $f(b)$ have opposite signs, the equation $f(x) = 0$ has at least one real root between a and b; in fact, an odd number of such roots, if an m-fold root is counted as m roots.*

The only argument* given here is one based upon geometrical intuition. We are stating that, if the points

$$(a, f(a)), \quad (b, f(b))$$

lie on opposite sides of the x-axis, the graph of $y = f(x)$ crosses the x-axis once, or an odd number of times, between the vertical lines through these two points. Indeed, the part of the graph between these verticals is a continuous curve having one and only one point on each intermediate vertical line, since the function has a single value for each value of x. This would not follow for the graph of $y^2 = x$.

* An arithmetical proof based upon a refined theory of irrational numbers is given in Weber's *Lehrbuch der Algebra*, ed. 2, vol. 1, p. 123.

13. Sign of a Polynomial. Given a polynomial

$$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n \qquad (a_0 \neq 0)$$

with real coefficients, we can find a positive number P such that $f(x)$ has the same sign as a_0x^n when $x > P$. In fact,

$$f(x) = x^n(a_0 + \phi), \quad \phi = \frac{a_1}{x} + \frac{a_2}{x^2} + \cdots + \frac{a_n}{x^n}.$$

By the last result in § 11, the numerical value of ϕ is less than that of a_0 when $1/x$ is positive and less than a sufficiently small positive number, say $1/P$, and hence when $x > P$. Then $a_0 + \phi$ has the same sign as a_0, and hence $f(x)$ the same sign as a_0x^n.

The last result holds also when x is a negative number sufficiently large numerically. For, if we set $x = -X$, the former case shows that $f(-X)$ has the same sign as $(-1)^n a_0 X^n$ when X is a sufficiently large positive number.

We shall therefore say briefly that, for $x = +\infty$, $f(x)$ has the same sign as a_0; while, for $x = -\infty$, $f(x)$ has the same sign as a_0 if n is even, but the sign opposite to a_0 if n is odd.

EXERCISES

1. $x^3 + ax^2 + bx - 4 = 0$ has a positive real root [use $x = 0$ and $x = +\infty$].
2. $x^3 + ax^2 + bx + 4 = 0$ has a negative real root [use $x = 0$ and $x = -\infty$].
3. If $a_0 > 0$ and n is odd, $a_0x^n + \cdots + a_n = 0$ has a real root of sign opposite to the sign of a_n [use $x = -\infty, 0, +\infty$].
4. $x^4 + ax^3 + bx^2 + cx - 4 = 0$ has a positive and a negative root.
5. Any equation of even degree n in which the coefficient of x^n and the constant term are of opposite signs has a positive and a negative root.

14. The accuracy of a graph of $y = f(x)$ can often be tested and important conclusions drawn from it by use of the

Theorem. *No straight line crosses the graph of $y = f(x)$ in more than n points if the degree n of the polynomial $f(x)$ exceeds unity.*

A vertical line $x = c$ crosses it at the single point $(c, f(c))$. A non-vertical line is the graph of an equation $y = mx + b$ of the first degree, and the abscissas of the points of crossing are the roots of $mx + b = f(x)$. The proof may now be completed by using the next theorem.

15. Theorem. *An equation of degree n,*

$$f(x) \equiv a_0x^n + a_1x^{n-1} + \cdots + a_n = 0 \qquad (a_0 \neq 0),$$

cannot have more than n distinct roots.

Suppose that it has the distinct roots $\alpha_1, \ldots, \alpha_n, \alpha$. By the Factor Theorem (§ 7), $x - \alpha_1$ is a factor of $f(x)$, so that

$$f(x) \equiv (x - \alpha_1)\, Q(x),$$

where $Q(x)$ is a polynomial of degree $n - 1$. Let $x = \alpha_2$. We see that $Q(\alpha_2) = 0$, so that as before

$$Q(x) \equiv (x - \alpha_2)\, Q_1(x), \quad f(x) \equiv (x - \alpha_1)(x - \alpha_2)\, Q_1(x).$$

Proceeding in this manner, we get

$$f(x) \equiv a_0(x - \alpha_1)(x - \alpha_2) \ldots (x - \alpha_n).$$

For the root α, the left member is zero and the right is not zero. Hence our supposition is false and the theorem true.

EXERCISES

1. The curve in Fig. 3, representing a cubic function, does not cross the x-axis at a second point further to the right, nor does the part starting from M' and running downwards to the left later ascend and cross the x-axis.
2. The curve in Fig. 2, representing a quartic function, has only the four crossings shown.
3. Form the cubic equation having the roots 0, 1, 2.
4. Form the quartic equation having the roots ± 1, ± 2.
5. If $a_0x^n + \cdots = 0$ has more than n distinct roots, each coefficient is zero. When would the theorem in § 14 fail if $n = 1$?
6. If two polynomials in x of degree n are equal for more than n distinct values of x, they are identical.
7. An equation of degree n cannot have more than n roots, a root of multiplicity m being counted as m roots.

16. Graphical Solution of a Quadratic Equation. If

$$x^2 - ax + b = 0 \tag{12}$$

has real coefficients and real roots, the roots may be constructed by the use of ruler and compasses, *i.e.*, by elementary geometry.

Draw a circle having as a diameter the line BQ joining the points $B = (0, 1)$ and $Q = (a, b)$; the abscissas ON and OM of the points of intersection of this circle with the x-axis are the roots of (12).

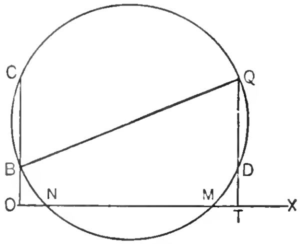


Fig. 9

The center of the circle is $(a/2, (b+1)/2)$. The square of BQ is $a^2 + (b-1)^2$. Hence the equation of the circle is

$$\left(x - \frac{a}{2}\right)^2 + \left(y - \frac{b+1}{2}\right)^2 = \left(\frac{a}{2}\right)^2 + \left(\frac{b-1}{2}\right)^2.$$

Setting $y = 0$, we get (12).

If we do not insist upon a solution by ruler and compasses, we may plot the parabola $y = x^2$ and draw the straight line $y = ax - b$; if these intersect, the abscissas of the points of intersection are the real roots of (12).

17. The method last used enables us to solve graphically

$$x^3 - ax + b = 0.$$

We have merely to employ the abscissas of the intersections of the graph (Fig. 4) of $y = x^3$ with $y = ax - b$. For the quartic equation

$$z^4 + Az^2 + Bz + C = 0, \; A > 0,$$

set $z = x\sqrt{A}$; we get

$$x^4 + x^2 - ax + b = 0.$$

We now employ the graphs of $y = x^4 + x^2$, $y = ax - b$.

EXERCISES

Solve by each of the two methods

1. $x^2 - 5x + 4 = 0$. 2. $x^2 + 5x + 4 = 0$. 3. $x^2 + 5x - 4 = 0$.
4. $x^2 - 5x - 4 = 0$. 5. $x^2 - 4x + 4 = 0$. 6. $x^2 - 3x + 4 = 0$.

Solve graphically the cubic equations

7. $x^3 - 3x + 1 = 0$. 8. $x^3 + 2x - 4 = 0$. 9. $x^3 - 7x + 7 = 0$.

10. Find graphically the cube roots of 20, -20, 200.

11. State in the language of elementary geometry the construction of Fig. 9 and prove that $OC = TQ = b$, $TD = OB = 1$, chord BN = chord DM, $ON = MT$, $ON + OM = a$, $ON \cdot OM = OC \cdot OB = b$. Why are OM and ON the roots of (12)?

12. Any reduced cubic equation $x^3 = px + q$ can be solved by use of a fixed parabola $x^2 = y$ and the circle $x^2 + y^2 = qx + (p+1)y$. (Descartes.)

13. $x^4 = px^2 + qx + r$ can be solved by use of a fixed parabola $x^2 = y$ and the circle $x^2 + y^2 = qx + (p+1)y + r$. (Descartes.)

14. Solve the cubics in Exs. 7–9 by the method of Ex. 12.

15. Solve $x^4 = 25x^2 - 60x + 36$ by the method of Ex. 13.

18.† The approximate values of the real roots of a cubic equation

$$z^3 + pz + q = 0$$

may be found by a graphical method due to C. Runge.* We assign equidistant values to z. For each z, we have a linear equation in p and q which therefore represents a straight line when p and q are taken as rectangular coördinates. On a diagram showing these lines we may locate approximately the line (and hence the values of z) corresponding to assigned values of p and q. The method applies also to any equation involving two parameters linearly.

For the solution of a numerical cubic equation by means of the slide rule (and an account of the use of the latter), see pp. 43–48 of the book just cited.

* *Graphical Methods*, Columbia University Press, 1912, p. 59 (also, *Praxis der Gleichungen*, Leipzig, 1900, p. 156). Earlier by L. Lalanne, Comptes Rendus Acad. Sc. Paris, 81, 1875, p. 1186, p. 1243; 82, 1876, p. 1487; 87, 1878, p. 157, and in Notices réunies par le Ministère des travaux . . . exposition univ. Paris, 1878.

CHAPTER II

Complex Numbers

(For a briefer course, this chapter may be begun with § 5.)

1.† Vectors from a Fixed Origin O. A directed segment of a straight line is called a *vector*. We shall employ only vectors from a fixed initial point O.

The *sum* of two vectors OA and OC is defined to be the vector OS, where S is the fourth vertex of the parallelogram having the lines OA and OC as two sides. In case A coincides with O, the vector OA is said to be zero; then $OS = OC$.

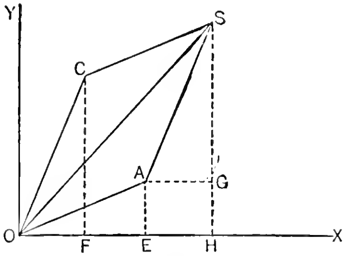


Fig. 10

A force of given magnitude and given direction is conveniently represented by a vector. By a fundamental principle of mechanics, two forces, represented by the vectors OA and OC, have as their resultant a force represented by the vector OS, as in Fig. 10. Thus if two forces are represented by two vectors, their resultant is represented by the sum of the vectors.

When referred to rectangular axes OX and OY, let the point A have the coördinates $OE = a$, $EA = b$, and the point C the coördinates $OF = c$, $FC = d$. Draw AG parallel to OX and SGH perpendicular to OX. Since triangles OFC and AGS are equal, $AG = c$, $GS = d$. Hence the coördinates of the point S are $OH = a + c$ and $HS = b + d$. *The sum of the vectors from O to the points (a, b) and (c, d) is the vector from O to the point $(a + c, b + d)$, whose coördinates are the sums of the corresponding coördinates of the two points.*

Subtraction of vectors is defined as the operation inverse to addition of vectors. If OA and OS are given vectors, the vector OC for which $OA + OC = OS$ is denoted by $OS - OA$, and is determined by the side OC of the parallelogram with the diagonal OS and side OA.

2.† Multiplication of Vectors. Let A be a point $\{r, \theta\}$ with the *polar coördinates* r, θ. Then r is the positive number giving the length of the line OA, while θ is the measure of the angle XOA when measured counter-clockwise from OX, as in Trigonometry. Let C be the point $\{r', \theta'\}$ with the polar coördinates r', θ'.

The product $OA \cdot OC$ of the vectors from O to $A = \{r, \theta\}$ and to $C = \{r', \theta'\}$ is defined to be the vector from O to $P = \{rr', \theta + \theta'\}$.

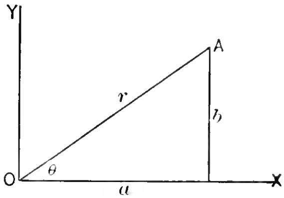


Fig. 11

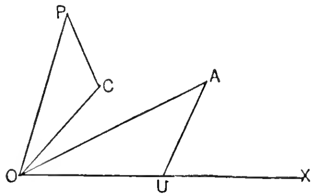


Fig. 12

To construct this product geometrically, let U be the point on the x-axis one unit to the right of O. Let the triangle OCP be constructed similar to triangle OUA, such that corresponding sides are OC and OU, CP and UA, OP and OA, and such that the vertices O, C, P are in the same order (clockwise or counter-clockwise) as the corresponding vertices O, U, A. Then $OP : r' = r : 1$, so that the length of OP is rr'. The angle XOP, measured counter-clockwise from OX, equals $\theta + \theta'$, and may exceed four right angles. Hence the product of the vectors OA and OC is the vector OP.

If $OC = OU$, then $OP = OA$, and $OU \cdot OA = OA$. Hence vector OU plays the rôle of unity in the multiplication of vectors.

Division of vectors is defined as the operation inverse to multiplication of vectors. If OA and OP are given vectors, the vector OC for which $OA \cdot OC = OP$ is denoted by OP/OA. If $A = \{r, \theta\}$ and $P = \{r_1, \theta_1\}$ then $C = \{r_1/r, \theta_1 - \theta\}$. Division except by zero is therefore always possible and unique.

EXERCISES †

1.† Vector addition is associative: $(OA + OC) + OL = OA + (OC + OL)$.

2.† Vector multiplication is associative: $(OA \cdot OC) \cdot OL = OA \cdot (OC \cdot OL)$.

3.† Draw the figure corresponding to Fig. 12, when OA is in the third quadrant and OC in the first quadrant.

3.† Symbol for Vectors from O. We consider only vectors starting from the fixed point O. Such a vector OA is uniquely determined by its terminal point $A = (a, b)$ and hence by the Cartesian coördinates a, b of the point A referred to fixed rectangular axes OX and OY. We may therefore denote the vector OA by the symbol $[a, b]$. Then

(1) $$[a, b] = [c, d] \text{ if and only if } a = c,\ b = d.$$

By the definition of addition and subtraction of vectors (§ 1),

(2) $$[a, b] + [c, d] = [a + c, b + d],$$

(3) $$[a, b] - [c, d] = [a - c, b - d].$$

As our definition of the product of two vectors was made in terms of polar coördinates, we must now express the product in terms of Cartesian coördinates. By Fig. 11, we have

$$a = r \cos \theta, \qquad b = r \sin \theta.$$

Similarly, if the point (c, d) has the polar coördinates r', θ',

$$c = r' \cos \theta', \quad d = r' \sin \theta'.$$

Hence the definition (§ 2) of the product of two vectors gives

$$[a, b]\,[c, d] = [rr' \cos (\theta + \theta'),\ rr' \sin (\theta + \theta')],$$

the final numbers being the Cartesian coördinates of the point with the polar coördinates rr' and $\theta + \theta'$. But

$$rr' \cos (\theta + \theta') = rr' (\cos \theta \cos \theta' - \sin \theta \sin \theta') = ac - bd,$$

$$rr' \sin (\theta + \theta') = rr' (\sin \theta \cos \theta' + \cos \theta \sin \theta') = bc + ad.$$

Hence, finally,

(4) $$[a, b]\,[c, d] = [ac - bd,\ ad + bc].$$

Given a, b, e, f, we can find solutions c, d of the equations

$$ac - bd = e, \quad ad + bc = f,$$

provided $a^2 + b^2 \neq 0$, viz., a and b are not both zero. Then

$$[a, b]\,[c, d] = [e, f]$$

determines $[c, d]$, its expression being

(5) $$\frac{[e, f]}{[a, b]} = \left[\frac{ae + bf}{a^2 + b^2},\quad \frac{af - be}{a^2 + b^2}\right].$$

Hence division, except by the zero vector $[0, 0]$, is always possible and unique.

4.† Introduction of Complex Numbers. Giving up the concrete interpretation in § 3 of the symbol $[x, y]$ as the vector from the origin to the point (x, y), we shall now think abstractly of a system of elements $[x, y]$ each determined by two real numbers x, y, and such that the system contains an element corresponding to any pair of real numbers. While the present abstract discussion is logically independent of the earlier exposition of vectors, yet we shall be guided in our present choice of definitions of addition, multiplication, etc., of our abstract symbols $[x, y]$ by the desire that the vector system shall furnish us a concrete representation of the present abstract system. Accordingly, we define equality, addition, subtraction, multiplication and division of two abstract elements $[x, y]$ by formulas (1)–(5). In particular, we have

$$[a, 0] \pm [c, 0] = [a \pm c, 0],$$

$$[a, 0]\,[c, 0] = [ac, 0], \quad \frac{[c, 0]}{[a, 0]} = \left[\frac{c}{a}, 0\right],$$

provided $a \neq 0$ in the last relation. Hence the elements $[x, 0]$ combine under our addition, multiplication, etc., exactly as the real numbers x combine under ordinary addition, multiplication, etc. We shall therefore introduce no contradiction if we now impose upon our abstract system of elements $[x, y]$, subject to relations (1)–(5), the further condition that the element $[x, 0]$ shall be the real number x. Then, by (4),

$$[0, 1]\,[0, 1] = [-1, 0] = -1.$$

We write i for $[0, 1]$. Hence $i^2 = -1$. Then

$$[x, y] = [x, 0] + [0, y] = x + [y, 0]\,[0, 1] = x + yi.$$

The resulting symbol $x + yi$ is called a *complex number*. For $y = 0$, it reduces to the real number x. For $y \neq 0$, it is also called an *imaginary number*. The latter is not to be thought of as unreal in the sense that its use is illogical. On the contrary, $x + yi$ is a convenient analytic representation of the vector from the origin to the point (x, y), and the sum, product, etc., defined above, of two such complex numbers then represent those simple combinations of the two corresponding vectors (§§ 1, 2) which are constantly used in the applications of vectors in mechanics and physics. Since these vectors from O are uniquely determined by their terminal points, we obtain a representation (§ 8) of complex numbers by points

in a plane, a representation of great importance in mathematics and its applications.

If in (1)–(5), we replace the symbol $[a, b]$ by $a + bi$, etc., we obtain the formulas given in § 5.

5. Formal Algebraic Definition of Complex Numbers. The equation $x^2 = -4$ has no real root, but is said to have the two imaginary roots $\sqrt{-4}$ and $-\sqrt{-4}$. We shall denote these roots by $2\,i$ and $-2\,i$, agreeing that i is a definite number for which $i^2 = -1$. Similarly, we shall write $\sqrt{3}\,i$ in preference to $\sqrt{-3}$. If p is positive, $\sqrt{p}$ is used to denote the positive square root of p.

If a and b are any two real numbers, $a + bi$ is called a *complex number* and $a - bi$ its *conjugate*. Two complex numbers $a + bi$ and $c + di$ are called equal if and only if $a = c$, $b = d$. Thus $a + bi = 0$ if and only if $a = b = 0$.

Addition of complex numbers is defined by

$$(a + bi) + (c + di) = (a + c) + (b + d)i.$$

The inverse operation, called subtraction, consists in finding a complex number z such that $(c + di) + z = a + bi$. In notation and value, z is

$$(a + bi) - (c + di) = (a - c) + (b - d)i.$$

Multiplication is defined by

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i,$$

and hence is performed as in formal algebra with a subsequent reduction by use of $i^2 = -1$. If we replace b by $-b$ and d by $-d$, the right member is replaced by its conjugate. Hence the product of the conjugates of two complex members equals the conjugate of their product.

Division is defined as the operation inverse to multiplication, and consists in finding a complex number q such that $(a + bi)q = e + fi$. Multiplying each member by $a - bi$, we find that q is, in notation and value,

$$\frac{e + fi}{a + bi} = \frac{(e + fi)(a - bi)}{a^2 + b^2} = \frac{ae + bf}{a^2 + b^2} + \frac{af - be}{a^2 + b^2}\,i.$$

Since $a^2 + b^2 = 0$ implies $a = b = 0$ when a and b are real, division except by zero is possible and unique.

6. The Cube Roots of Unity. The roots of $x^3 = 1$ are unity and the numbers for which

$$\frac{x^3 - 1}{x - 1} \equiv x^2 + x + 1 = 0, \quad (x + \tfrac{1}{2})^2 = -\tfrac{3}{4}, \quad x + \tfrac{1}{2} = \pm \tfrac{1}{2}\sqrt{3}\,i.$$

Hence the three cube roots of unity are 1 and

$$\omega = -\tfrac{1}{2} + \tfrac{1}{2}\sqrt{3}\,i, \quad \omega' = -\tfrac{1}{2} - \tfrac{1}{2}\sqrt{3}\,i.$$

EXERCISES

1. Verify that $\omega' = \omega^2$, $\omega\omega' = 1$, $\omega^2 + \omega + 1 = 0$, $\omega^3 = 1$.
2. The sum and product of two conjugate complex numbers are real.
3. Express as complex numbers

$$\frac{3 + 5\,i}{2 - 3\,i}, \quad \frac{a + bi}{a - bi}, \quad \frac{3 + \sqrt{-5}}{2 + \sqrt{-1}}.$$

4. If x, y, z are any complex numbers,

$$x + y = y + x, \quad (x + y) + z = x + (y + z),$$

$$xy = yx, \quad (xy)z = x(yz), \quad x(y + z) = xy + xz.$$

What is the name of the property indicated by each equation?

5. If the product of two complex numbers is zero, one of them is zero.

6.† Deduce the laws in § 5 from those in § 4.

7. Square Roots of $a + bi$ found Algebraically. Given the real numbers a and b, $b \neq 0$, we seek real numbers x and y such that

$$a + bi = (x + yi)^2 \equiv x^2 - y^2 + 2\,xyi.$$

Thus

$$x^2 - y^2 = a, \quad 2\,xy = b,$$

$$(x^2 + y^2)^2 = (x^2 - y^2)^2 + 4\,x^2y^2 = a^2 + b^2.$$

Since x and y are to be real and hence $x^2 + y^2$ positive,

$$x^2 + y^2 = \sqrt{a^2 + b^2},$$

the positive square root being the one taken. Combining this equation with $x^2 - y^2 = a$, we get

$$x^2 = \frac{\sqrt{a^2 + b^2} + a}{2}, \quad y^2 = \frac{\sqrt{a^2 + b^2} - a}{2}.$$

Since these expressions are positive, real values of x and y may be found. The two pairs x, y for which $2xy = b$ give the desired two complex numbers $x + yi$.

It is not possible to find the cube roots of a general complex number by a similar algebraic process (Ch. III, § 6).

EXERCISES

Express as complex numbers the square roots of

1. $-7 + 24i$.
2. $-11 + 60i$.
3. $5 - 12i$.
4. $4cd + (2c^2 - 2d^2)i$.
5. $c^2 - d^2 - 2\sqrt{-c^2d^2}$.

8. Geometrical Representation of Complex Numbers. Using rectangular axes of coördinates, we represent* $a + bi$ by the point $A = (a, b)$. The positive number $r = \sqrt{a^2 + b^2}$ giving the length of OA is called the *modulus* (or *absolute value*) of $a + bi$ (Fig. 11). The angle $\theta = XOA$, measured counter-clockwise from OX, is called the *amplitude* (or *argument*) of $a + bi$. Thus

$$a + bi = r(\cos\theta + i\sin\theta). \tag{6}$$

The second member is called the *trigonometric form* of $a + bi$.

If $c + di$ is represented by the point C, then the sum of $a + bi$ and $c + di$ is the complex number represented by the point S (Fig. 10) determined by the parallelogram $OASC$. Since $OS \leqq OA + AS$, *the modulus of the sum of two complex numbers is equal to or less than the sum of their moduli.*

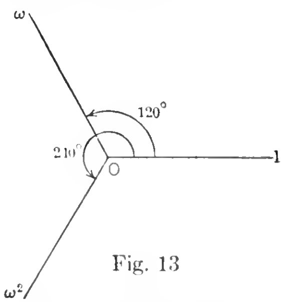


Fig. 13

For example, the cube roots of unity are 1 and

$$\omega = -\tfrac{1}{2} + \tfrac{1}{2}\sqrt{3}\,i$$
$$= \cos 120° + i \sin 120°,$$

$$\omega^2 = -\tfrac{1}{2} - \tfrac{1}{2}\sqrt{3}\,i$$
$$= \cos 240° + i \sin 240°,$$

and are respresented by the points marked 1, ω, ω^2 in Fig. 13. They form

* It will be obvious to the reader who has not omitted §§ 1–4 that the present representation is essentially equivalent to the representation of $a + bi$ by the vector from O to the point (a, b).

the vertices of an equilateral triangle inscribed in a circle of unit radius and center at the origin O.

9. The product of the complex number (6) by $r'(\cos\alpha + i\sin\alpha)$ is

$$rr'\,[\cos(\theta+\alpha) + i\sin(\theta+\alpha)],$$

since

$$(\cos\theta + i\sin\theta)(\cos\alpha + i\sin\alpha) = \cos(\theta+\alpha) + i\sin(\theta+\alpha). \tag{7}$$

The latter follows from

$$\cos\theta\cos\alpha - \sin\theta\sin\alpha = \cos(\theta+\alpha),$$

$$\cos\theta\sin\alpha + \sin\theta\cos\alpha = \sin(\theta+\alpha).$$

Hence *the modulus of the product of two complex numbers equals the product of their moduli, and the amplitude of the product equals the sum of their amplitudes.*

The product may be found geometrically as in Fig. 12.

For the special case $\alpha = \theta$, (7) becomes

$$(\cos\theta + i\sin\theta)^2 = \cos 2\theta + i\sin 2\theta.$$

This is the case $n = 2$ of formula (8). In particular, we see why the amplitude of ω^2 is 240° when that of ω is 120° (end of § 8).

10. De Moivre's Theorem. *If n is any positive integer,*

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta. \tag{8}$$

This relation is an identity if $n = 1$ and was seen to hold if $n = 2$. To proceed by mathematical induction, let it be true if $n = m$. Using (7) for $\alpha = m\theta$, we then have

$$(\cos\theta + i\sin\theta)^{m+1} = (\cos\theta + i\sin\theta)(\cos\theta + i\sin\theta)^m$$

$$= (\cos\theta + i\sin\theta)(\cos m\theta + i\sin m\theta) = \cos(m+1)\theta + i\sin(m+1)\theta.$$

Hence (8) is true also if $n = m + 1$. The induction is thus complete.

Since $\cos\theta + i\sin\theta$ represents the vector from the origin O to the point $\{1, \theta\}$, given in polar coördinates, its nth power represents (§ 2) the vector from O to the point $\{1, n\theta\}$ and hence is $\cos n\theta + i\sin n\theta$.

11. Cube Roots. To find the cube roots of a complex number, we first express it in the trigonometric form (6). For example,

$$4\sqrt{2} + 4\sqrt{2}\,i = 8\,(\cos 45° + i \sin 45°).$$

If it has a cube root of the form (6), then, by (8),

$$r^3\,(\cos 3\,\theta + i \sin 3\,\theta) = 8\,(\cos 45° + i \sin 45°).$$

Their moduli r^3 and 8 must be equal, so that the positive real number r equals 2. Since $3\,\theta$ and $45°$ have equal cosines and equal sines, they differ by an integral multiple of $360°$. Thus

$$\theta = 15° + k \cdot 120° \qquad (k \text{ an integer}).$$

Since in (6) we may replace θ by $\theta + 360°$ without changing $a + bi$, we obtain just three distinct cube roots (given by $k = 0, 1, 2$):

$$2\,(\cos 15° + i \sin 15°),\ 2\,(\cos 135° + i \sin 135°),\ 2\,(\cos 255° + i \sin 255°).$$

EXERCISES

1. Verify that the last two numbers equal the products of the first number by ω and ω^2, given at the end of § 8.
2. Find the three cube roots of -27; those of $-i$.
3. Find the three cube roots of $-\frac{1}{2} + \frac{1}{2}\sqrt{3}\,i$.

12. nth Roots. Let ρ be a positive real number. As illustrated in § 11, it is evident that the nth roots of $\rho\,(\cos A + i \sin A)$ are the products of the nth roots of $\cos A + i \sin A$ by the positive real nth root of ρ. Let an nth root of $\cos A + i \sin A$ be of the form (6). Then, by (8),

$$r^n(\cos n\theta + i \sin n\theta) = \cos A + i \sin A.$$

Thus $r^n = 1$, $r = 1$, and $n\theta = A + k \cdot 360°$, where k is an integer. Thus n distinct nth roots of $\cos A + i \sin A$ are given by

$$\cos \frac{A + k \cdot 360°}{n} + i \sin \frac{A + k \cdot 360°}{n} \qquad (k = 0, 1, \ldots, n-1), \tag{9}$$

whereas $k = n$ gives the same root as $k = 0$, and $k = n + 1$ the same root as $k = 1$, etc. Hence any number $\neq 0$ has exactly n distinct nth complex roots.

EXERCISES

1. Find the five fifth roots of -1.
2. Find the nine ninth roots of 1. Which are roots of $x^3 = 1$?
3. Simplify the trigonometric forms of the four fourth roots of unity. Check the result by factoring $x^4 - 1$.

13. Roots of Unity. By (9) the n distinct nth roots of unity are

$$(10) \qquad \cos\frac{2\,k\pi}{n} + i\sin\frac{2\,k\pi}{n} \qquad (k = 0,\ 1,\ \ldots,\ n-1),$$

where now the angles are measured in radians (an angle of 180 degrees equals π radians, where $\pi = 3.1416$, approximately). For $k = 0$, (10) reduces to 1, which is an evident nth root of unity. For $k = 1$, (10) is

$$(11) \qquad r = \cos\frac{2\,\pi}{n} + i\sin\frac{2\,\pi}{n}.$$

By De Moivre's Theorem (§ 10), the general number (10) equals the kth power of r. Hence the n distinct nth roots of unity are

$$(12) \qquad r,\ r^2,\ r^3,\ \ldots,\ r^{n-1},\ r^n = 1.$$

The n complex numbers (10), and therefore the numbers (12), are represented geometrically by the vertices of a regular polygon of n sides inscribed in the circle of radius unity and center at the origin with one vertex on the x-axis (Fig. 14).

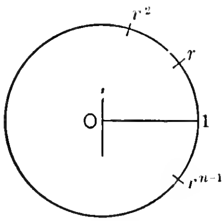


Fig. 14

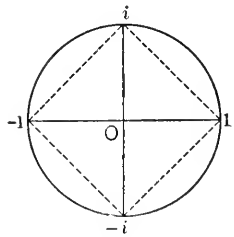


Fig. 15

For $n = 3$, the numbers (12) are ω, ω^2, 1, shown in Fig. 13.

For $n = 4$, we have $r = \cos\pi/2 + i\sin\pi/2 = i$. The fourth roots of unity (12) are i, $i^2 = -1$, $i^3 = -i$, $i^4 = 1$. These are represented by the vertices of a square inscribed in a circle of radius unity (Fig. 15).

EXERCISES

1. For $n = 6$, $r = -\omega^2$. The sixth roots of unity are therefore the three cube roots of unity and their negatives. Check by factoring $x^6 - 1$.

2. From the point representing $a + bi$ how do you obtain that representing $-(a + bi)$? Hence derive from Fig. 13 and Ex. 1 the points representing the six sixth roots of unity.

3. Which powers of a ninth root (11) of unity are cube roots of unity?

14. Primitive nth Roots of Unity. An nth root of unity is called *primitive* if no power of it, with a positive integral exponent less than n, equals unity. Since only the last one of the numbers (12) equals unity, the number r, given by (11), is a primitive nth root of unity.

For $n = 4$, both i and $-i$ are primitive fourth roots of unity, while 1 and -1 are not. Just as $i^2 = -1$ and $i^4 = +1$ are not primitive fourth roots of unity, so r^k is not a primitive nth root of unity if k and n have a common divisor d $(d > 1)$. Indeed,

$$(r^k)^{\frac{n}{d}} = (r^n)^{\frac{k}{d}} = 1,$$

whereas n/d is a positive integer less than n. But if k and n are relatively prime, *i.e.*, have no common divisor exceeding unity, r^k is a primitive nth root of unity. To prove this, we must show that $(r^k)^l \neq 1$ if l is a positive integer less than n. Now, by De Moivre's Theorem,

$$r^{kl} = \cos\frac{2\,kl\pi}{n} + i\sin\frac{2\,kl\pi}{n}.$$

If this were unity, $2\,kl\pi/n$ would be a multiple of $2\,\pi$, and hence kl a multiple of n. Since k is relatively prime to n, the second factor l would be a multiple of n, whereas $0 < l < n$. Hence *the primitive nth roots of unity are those of the numbers* (12) *whose exponents are relatively prime to n.*

EXERCISES

1. The primitive cube roots of unity are ω and ω^2.

2. For r given by (11), the primitive nth roots of unity are (i) for $n = 6$, r, r^5; (ii) for $n = 12$, r, r^5, r^7, r^{11}.

3. For n a prime, any nth root of unity, other than 1, is primitive.

4. If r is a primitive 15th root of unity, r^3, r^6, r^9, r^{12} are the primitive 5th roots of unity, and r^5, r^{10} are the primitive cube roots of unity. Show that their 8 products by pairs give all of the primitive 15th roots of unity.

5. If n is the product of two primes p and q, there are exactly $(p-1)(q-1)$ primitive nth roots of unity.

6. If ρ is any primitive nth root of unity, $\rho, \rho^2, \rho^3, \ldots, \rho^n$ are distinct and give all of the nth roots of unity. Of these, ρ^k is a primitive nth root of unity if and only if k is relatively prime to n.

15. Imaginary Roots Occur in Pairs. The roots of $x^2 + 2\,cx + d = 0$ are

$$-c + \sqrt{c^2 - d}, \qquad -c - \sqrt{c^2 - d}. \tag{13}$$

If c and d are real, these roots are both real or are conjugate imaginaries. The latter case illustrates the following

Theorem. *If a and b are real numbers, $b \neq 0$, and if $a + bi$ is a root of an equation with real coefficients, then $a - bi$ is a root.*

Let the equation be $f(x) = 0$. Divide $f(x)$ by

$$(x - a)^2 + b^2 \equiv (x - a - bi)(x - a + bi) \tag{14}$$

until we reach a remainder $rx + s$ of degree less than the degree of the divisor in x. Evidently r and s are real. If the quotient is $Q(x)$, we have

$$f(x) \equiv Q(x) \{(x - a)^2 + b^2\} + rx + s,$$

identically in x (Ex. 6, p. 15). Let $x = a + bi$. Since this is a root of $f(x) = 0$, we see that

$$0 = r(a + bi) + s, \quad 0 = ra + s, \quad 0 = rb.$$

Since $b \neq 0$, we have $r = 0$ and then $s = 0$. Thus $f(x)$ has the factor (14), so that $f(x) = 0$ has the root $a - bi$.

16.† Generalization of the theorem in § 15. The sum of the roots (13) of $x^2 + 2\,cx + d = 0$ equals the negative of the coefficient $2\,c$ of x, and their product equals the constant term d. It follows that $2 + i$ and -2 are the roots of

$$z^2 - iz - 4 - 2\,i = 0,$$

and that $2 - i$ and -2 are the roots of

$$z^2 + iz - 4 + 2\,i = 0.$$

We have here an illustration of the following

Theorem. *If a and b are real numbers and if $a + bi$ is a root of $f(z) = 0$, then $a - bi$ is a root of $g(z) = 0$, where $g(z)$ is obtained from the polynomial $f(z)$ by replacing each coefficient $c + di$ by its conjugate $c - di$.*

Consider any term $(c + di)z^k$ of $f(z)$. Replace z by $x + yi$, where x and y are real. The term

$$(c + di)(x + yi)^k$$

of $f(x + yi)$ has as its conjugate imaginary the product

$$(c - di)(x - yi)^k$$

of the conjugates of the factors of that term (§ 5). But the new product is a term of $g(x - yi)$. Hence the latter is the conjugate $A - Bi$ of $f(x + yi) \equiv A + Bi$, where A and B are polynomials in x and y with real coefficients.

Take $x = a$, $y = b$. Then $A = B = 0$ by hypothesis. Hence $g(a - bi) = 0$.

EXERCISES

1.† The theorem in § 15 is a corollary to that in § 16.

2. Solve $x^3 - 3x^2 - 6x - 20 = 0$, with the root $-1 + \sqrt{-3}$.

3. Solve $x^4 - 4x^3 + 5x^2 - 2x - 2 = 0$, with the root $1 - i$.

4. Find the cubic equation with real coefficients two of whose roots are 1 and $3 + 2i$.

5.† Given that $x^3 + (1 - i)x^2 + 1 = 0$ has the root i, find a cubic equation with the root $-i$. Form an equation with real coefficients whose roots include the roots of these two cubic equations.

6. If an equation with *rational* coefficients has a root $a + \sqrt{b}$, where a and b are rational, but $\sqrt{b}$ is irrational, it has the root $a - \sqrt{b}$. [Use the method of § 15.]

7. Solve $x^4 - 4x^3 + 4x - 1 = 0$, with the root $2 + \sqrt{3}$.

8. Solve $x^3 - (4 + \sqrt{3})x^2 + (5 + 4\sqrt{3})x - 5\sqrt{3} = 0$, with the root $\sqrt{3}$.

9. Solve the equation in Ex. 8, given that it has the root $2 + i$.

10. What cubic equation with rational coefficients has the roots $\frac{1}{2}$, $\frac{1}{2} + \sqrt{2}$?

CHAPTER III

Algebraic and Trigonometric Solution of Cubic Equations

1. Reduced Cubic Equation. If in the general cubic equation

$$x^3 + bx^2 + cx + d = 0, \tag{1}$$

we set $x = y - b/3$, we obtain a reduced cubic equation

$$y^3 + py + q = 0, \tag{2}$$

where

$$p = c - \frac{b^2}{3}, \qquad q = d - \frac{bc}{3} + \frac{2\,b^3}{27}. \tag{3}$$

A geometrical interpretation of this process was given in Ex. 5, p. 10.

We shall find the roots y_1, y_2, y_3 of (2). Then the roots of (1) are

$$x_1 = y_1 - \frac{b}{3}, \quad x_2 = y_2 - \frac{b}{3}, \quad x_3 = y_3 - \frac{b}{3}. \tag{4}$$

2. Algebraic Solution of Cubic Equation (2). We shall employ a method essentially that given by Vieta * in 1591. We make the substitution

$$y = z - \frac{p}{3\,z} \tag{5}$$

in (2) and obtain

$$z^3 - \frac{p^3}{27\,z^3} + q = 0.$$

Multiplying each member by z^3, we get

$$z^6 + qz^3 - \frac{p^3}{27} = 0. \tag{6}$$

Solving this as a quadratic equation for z^3, we obtain

$$z^3 = -\frac{q}{2} \pm \sqrt{R}, \quad R = \left(\frac{p}{3}\right)^3 + \left(\frac{q}{2}\right)^2. \tag{7}$$

* *Opera Math.*, IV, published by A. Anderson, Paris, 1615.

By Ch. II, § 11, any number has three cube roots, two of which are the products of the remaining one by

(8) $$\omega = -\tfrac{1}{2} + \tfrac{1}{2}\sqrt{3}\,i, \quad \omega^2 = -\tfrac{1}{2} - \tfrac{1}{2}\sqrt{3}\,i.$$

Since

$$\left(-\frac{q}{2} + \sqrt{R}\right)\left(-\frac{q}{2} - \sqrt{R}\right) = \left(-\frac{p}{3}\right)^3,$$

we can choose particular cube roots

(9) $$A = \sqrt[3]{-\frac{q}{2} + \sqrt{R}}, \quad B = \sqrt[3]{-\frac{q}{2} - \sqrt{R}},$$

such that $AB = -p/3$. Then the six values of z are

$$A,\ \omega A,\ \omega^2 A,\ B,\ \omega B,\ \omega^2 B.$$

These can be paired so that the product of the two in each pair is $-p/3$:

$$AB = -p/3, \quad \omega A \cdot \omega^2 B = -p/3, \quad \omega^2 A \cdot \omega B = -p/3.$$

Hence with any root z is paired a root equal to $-p/(3z)$. By (5), the sum of the two is a value of y. Thus the three roots of (2) are

(10) $$y_1 = A + B, \quad y_2 = \omega A + \omega^2 B, \quad y_3 = \omega^2 A + \omega B.$$

These are known as *Cardan's formulæ* for the roots of a reduced cubic equation (2). The expression $A + B$ for a root was first published by Cardan in his *Ars Magna*, 1545, although he had obtained it from Tartaglia under promise of secrecy.

EXERCISES

1. For $y^3 - 15y - 126 = 0$, $y = z + 5/z$ and
$$z^6 - 126z^3 + 125 = 0, \qquad z^3 = 1 \text{ or } 125, \qquad z = 1,\ \omega,\ \omega^2,\ 5,\ 5\omega,\ 5\omega^2.$$
The first three z's give the distinct y's: $6,\ \omega + 5\omega^2,\ \omega^2 + 5\omega$.

2. Solve $y^3 - 18y + 35 = 0$.
3. Solve $x^3 + 6x^2 + 3x + 18 = 0$.
4. Solve $y^3 - 2y + 4 = 0$.
5. Solve $28x^3 + 9x^2 - 1 = 0$.
6. Using $\omega^2 + \omega + 1 = 0$, show from (10) that
$$y_1 + y_2 + y_3 = 0, \qquad y_1y_2 + y_1y_3 + y_2y_3 = p, \qquad y_1y_2y_3 = -q.$$
7. By (3), (4) and Ex. 6, show that, for the roots of (1),
$$x_1 + x_2 + x_3 = -b, \qquad x_1x_2 + x_1x_3 + x_2x_3 = c, \qquad x_1x_2x_3 = -d.$$

3. Discriminant. By (10) and $\omega^3 = 1$,

$$y_1 - y_2 = (1-\omega)(A - \omega^2 B),$$
$$y_1 - y_3 = -\omega^2(1-\omega)(A - \omega B),$$
$$y_2 - y_3 = \omega(1-\omega)(A - B).$$

To form the product of these, note that $\omega^3 = 1$ and, by (8),

$$(1-\omega)^3 = 3(\omega^2 - \omega) = -3\sqrt{3}\, i.$$

Since the cube roots of unity are $1, \omega, \omega^2$, we have

$$x^3 - 1 \equiv (x-1)(x-\omega)(x-\omega^2),$$

identically in x. Taking $x = A/B$, we see that

$$A^3 - B^3 = (A-B)(A-\omega B)(A-\omega^2 B). \tag{11}$$

The left member equals $2\sqrt{R}$ by (9). Hence

$$(y_1-y_2)(y_1-y_3)(y_2-y_3) = 6\sqrt{3}\sqrt{R}\, i. \tag{12}$$

The product of the squares of the differences of the roots of any equation in which the coefficient of the highest power of the variable is unity shall be called the *discriminant* of the equation. Thus the discriminant is zero if and only if two roots are equal, and is positive if all the roots are real.

In view of (12) the discriminant Δ of the reduced cubic equation (2) has the value

$$\Delta = -108\, R = -4\, p^3 - 27\, q^2. \tag{13}$$

By (4), $x_1 - x_2 = y_1 - y_2$, etc. Hence the discriminant of the general cubic (1) equals the discriminant of the corresponding reduced cubic (2). By (3) and (13),

$$\Delta = 18\, bcd - 4\, b^3 d + b^2 c^2 - 4\, c^3 - 27\, d^2. \tag{14}$$

It is sometimes convenient to employ a cubic equation

$$ax^3 + bx^2 + cx + d = 0, \tag{15}$$

in which the coefficient of x^3 has not been made unity by division. The product P of the squares of the differences of its roots is evidently derived from (14) by replacing b, c, d by $b/a, c/a, d/a$. Thus

$$a^4 P = 18\, abcd - 4\, b^3 d + b^2 c^2 - 4\, ac^3 - 27\, a^2 d^2. \tag{16}$$

This expression (and not P itself) is called the discriminant * of (15).

* Some writers define $-\frac{1}{27}\, a^4 P$ to be the discriminant of (15) and hence $-\frac{1}{27}\,\Delta$ as that of (1). On this point see Ch. IV, § 4.

4. Theorem. *A cubic equation with real coefficients has three distinct real roots, a single real root, or at least two equal real roots, according as its discriminant is positive, negative or zero.*

It suffices to prove the theorem for a reduced cubic equation (2) in which p and q are real. First, let $\Delta \leqq 0$. By (13), $R \geqq 0$. Using (8), we find that the roots (10) are

$$A + B, \quad -\tfrac{1}{2}(A + B) \pm \tfrac{1}{2}(A - B)\sqrt{3}\, i. \tag{17}$$

But A and B, in (9), may now be taken to be real, since $R \geqq 0$.

If $R > 0$, $A \neq B$ and $A + B$ is the only real root. If $R = 0$, then $A = B$ and the roots are real and at least two are equal.

Next, let $\Delta > 0$, so that $R < 0$. Since $-\frac{1}{2}q + \sqrt{R}$ is an imaginary number it has (Ch. II, § 11) a cube root of the form $A = \alpha + \beta i$, where α and β are real and $\beta \neq 0$. Then (Ch. II, § 16) $B = \alpha - \beta i$ is a cube root of $-\frac{1}{2}q - \sqrt{R}$. For these cube roots, the product AB is real and hence equals $-p/3$, as required in § 2. Hence

$$y_1 = 2\alpha, \quad y_2 = -\alpha - \beta\sqrt{3}, \quad y_3 = -\alpha + \beta\sqrt{3}.$$

These real roots are distinct since $\Delta \neq 0$.

EXERCISES

Find by means of Δ the number of real roots of

1. $y^3 - 15y + 4 = 0$. 2. $y^3 - 27y + 54 = 0$. 3. $x^3 + 4x^2 - 11x + 6 = 0$.

4. Using $\Delta = (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2$, show that, if x_1 and x_2 are conjugate imaginaries and hence x_3 real, $\Delta < 0$; if the x's are all real and distinct, $\Delta > 0$. Deduce the theorem of § 4.

5. Deduce the same theorem from Ch. I, § 9.

5. Irreducible Case. When the roots of a cubic equation are all real and distinct, R is negative (§ 4), so that Cardan's formulæ present their values in a form involving cube roots of imaginaries. This is called the irreducible case.* We shall derive modified formulæ suitable for numerical work. Since any complex number can be expressed in the trigonometric form, we can find r and θ such that

$$-\tfrac{1}{2}q + \sqrt{R} = r(\cos\theta + i\sin\theta). \tag{18}$$

* This term is not to be confused with "irreducible equation."

In fact, the conditions for this equality are

$$-\tfrac{1}{2}q = r\cos\theta, \quad R = -r^2\sin^2\theta.$$

Hence

$$r^2 = r^2(\cos^2\theta + \sin^2\theta) = \tfrac{1}{4}q^2 - R = \frac{-p^3}{27},$$

(19) $$r = \sqrt{\frac{-p^3}{27}}, \quad \cos\theta = \frac{-q}{2} \div \sqrt{\frac{-p^3}{27}}.$$

Since R is negative, p is negative and r is real. Since $R < 0$, the value (19) of $\cos\theta$ is numerically less than unity. Hence θ can be found from a table of cosines.

The complex number conjugate to (18) is

(20) $$-\tfrac{1}{2}q - \sqrt{R} = r(\cos\theta - i\sin\theta).$$

The cube roots of (18) and (20) are

$$\sqrt{\frac{-p}{3}}\left[\cos\frac{\theta + m\cdot 360°}{3} \pm i\sin\frac{\theta + m\cdot 360°}{3}\right] \qquad (m = 0, 1, 2).$$

For a fixed value of m the product of these two numbers is $-p/3$. Hence their sum is a root of our cubic equation. Thus *if R is negative, the three distinct real roots are*

(21) $$2\sqrt{\frac{-p}{3}}\cos\frac{\theta + m\cdot 360°}{3} \qquad (m = 0, 1, 2).$$

EXERCISES

1. Solve the cubics in Exs. 1, 2, page 34.
2. Solve $y^3 - 2y - 1 = 0$.
3. Solve $y^3 - 7y + 7 = 0$.
4. Find constants r and s such that

$$y^3 + py + q \equiv \frac{1}{r - s}\{r(y + s)^3 - s(y + r)^3\}$$

identically in y. Hence solve the reduced cubic equation.

6.† Algebraic Discussion of the Irreducible Case. Avoiding the use of trigonometric functions, we shall attempt to find algebraically an exact cube root $x + yi$ of $a + bi$, where a and b are given real numbers, $b \neq 0$. We desire real numbers x and y such that

$$(x + yi)^3 = a + bi,$$

whence $$x^3 - 3xy^2 = a, \quad 3x^2y - y^3 = b.$$

Thus $y \neq 0$ and we may therefore set $x = sy$. Hence

$$(s^3 - 3s)y^3 = a, \quad (3s^2 - 1)y^3 = b$$

Eliminating y^3, we get

$$s^3 - \frac{3a}{b}s^2 - 3s + \frac{a}{b} = 0.$$

Set $s = t + a/b$. We obtain the reduced cubic equation

$$t^3 - 3kt - 2\frac{a}{b}k = 0 \qquad \left(k = \frac{a^2}{b^2} + 1\right).$$

The R of (7) is here $-k^2$. Thus Cardan's formulæ for the roots t involve

$$A = \sqrt[3]{\frac{a}{b}k + ki} = \sqrt[3]{\frac{k}{b}} \cdot \sqrt[3]{a + bi}.$$

While the first factor is the cube root of a real number, the second is exactly the cube root which we started out to find.

Hence this algebraic process in conjunction with that in § 2 fails to give us the real roots of our cubic equation. Conceivably other algebraic processes would succeed; but it can be proved * rigorously that a cubic equation with rational coefficients having no rational root, but having three real roots, cannot be solved in terms of real radicals only. Hence there does not exist an algebraic process for finding the real values of the roots in the irreducible case.

A cube root of a general complex number cannot be expressed in the form $x + yi$, where x and y involve only real radicals. For, if so, Cardan's formulæ could be simplified so as to express the roots of any cubic equation in terms of real radicals only.

7.† Trigonometric Solution of a Cubic Equation with $\Delta > 0$. In the irreducible case we may avoid Cardan's formulæ and the simplifications in § 5. The same final results are now obtained by a direct solution based upon the well-known trigonometric identity

$$\cos 3x = 4\cos^3 x - 3\cos x.$$

* H. Weber and J. Wellstein, *Encyklopädie der Elementar-Mathematik*, I, ed. 1, p. 325; ed. 2, p. 373; ed. 3, p. 364.

This may be written in the form

$$z^3 - \tfrac{3}{4} z - \tfrac{1}{4} \cos 3x = 0 \qquad (z = \cos x).$$

To transform cubic (2) into this one, set $y = nz$. Thus

$$z^3 + \frac{p}{n^2} z + \frac{q}{n^3} = 0.$$

The two cubic equations are identical if

$$n = \sqrt{\frac{-4p}{3}}, \quad \cos 3x = \frac{-q}{2} \div \sqrt{\frac{-p^3}{27}}.$$

Since $R < 0$, $p < 0$ and the value of $\cos 3x$ is real and numerically < 1. Hence we can find $3x$ from a table of cosines. The three values of z are then

$$\cos x, \quad \cos(x + 120°), \quad \cos(x + 240°).$$

Multiplying these by n, we get the three roots y.

Example. For $y^3 - 2y - 1 = 0$, we have

$$n^2 = 8/3, \quad \cos 3x = \sqrt{27/32}, \quad 3x = 23° \, 17' \, 0'',$$

$$\cos x = 0.99084, \quad \cos(x + 120°) = -0.61237, \quad \cos(x + 240°) = -0.37847,$$

$$y = 1.61804, \ -1, \ -0.61804.$$

EXERCISES†

Solve by the last method

1. $y^3 - 7y + 7 = 0$.
2. $x^3 + 3x^2 - 2x - 5 = 0$.
3. $x^3 + x^2 - 2x - 1 = 0$.
4. $x^3 + 4x^2 - 7 = 0$.
5. The cubic for t in § 6 has three real roots; in just three of the nine sets of solutions x, y, both are real.

CHAPTER IV

Algebraic Solution of Quartic Equations

1. Ferrari's Method. Writing the quartic equation

$$(1) \qquad x^4 + bx^3 + cx^2 + dx + e = 0$$

in the equivalent form

$$(x^2 + \tfrac{1}{2} bx)^2 = (\tfrac{1}{4} b^2 - c)x^2 - dx - e$$

and adding $(x^2 + \frac{1}{2} bx)y + \frac{1}{4} y^2$ to each member, we get

$$(2) \qquad (x^2 + \tfrac{1}{2} bx + \tfrac{1}{2} y)^2 = (\tfrac{1}{4} b^2 - c + y)x^2 + (\tfrac{1}{2} by - d)x + \tfrac{1}{4} y^2 - e.$$

We seek a value y_1 of y such that the second member of (2) shall be the square of a linear function of x. For brevity, write

$$(3) \qquad b^2 - 4c + 4y_1 = t^2.$$

We here assume that $t \neq 0$ (*cf.* Exs. 3, 4, p. 40). We therefore desire that

$$(4) \qquad \tfrac{1}{4} t^2x^2 + (\tfrac{1}{2} by_1 - d)x + \tfrac{1}{4} y_1^2 - e = \left(\tfrac{1}{2} tx + \frac{\frac{1}{2} by_1 - d}{t}\right)^2.$$

The condition for this is that the terms free of x be equal:

$$(5) \qquad \tfrac{1}{4} y_1^2 - e = \frac{(\frac{1}{2} by_1 - d)^2}{b^2 - 4c + 4y_1}.$$

Hence y_1 must be a root of the *resolvent cubic equation*

$$(6) \qquad y^3 - cy^2 + (bd - 4e)y - b^2e + 4ce - d^2 = 0.$$

After finding (Ch. III) a root y_1 of this cubic equation, we can easily get the roots of the quartic equation. In view of (2) and (4), each root of the quartic equation satisfies one of the quadratic equations

$$(7) \qquad \begin{cases} x^2 + \frac{1}{2}(b - t)x + \frac{1}{2} y_1 - (\frac{1}{2} by_1 - d)/t = 0, \\ x^2 + \frac{1}{2}(b + t)x + \frac{1}{2} y_1 + (\frac{1}{2} by_1 - d)/t = 0. \end{cases}$$

EXERCISES

1. For $x^4 + 2x^3 - 12x^2 - 10x + 3 = 0$, show that (6) becomes $y^3 + 12y^2 - 32y - 256 = 0$, with the root $y_1 = -4$, and that (7) then become

$$x^2 + 4x - 1 = 0, \quad x^2 - 2x - 3 = 0,$$

with the roots $-2 \pm \sqrt{5}$; 3, -1.

2. Solve $x^4 - 2x^3 - 7x^2 + 8x + 12 = 0$.
3. Solve $x^4 - 8x^3 + 9x^2 + 8x - 10 = 0$.

2. Relations between the Roots and Coefficients. Let x_1 and x_2 be the roots of the first quadratic equation (7), x_3 and x_4 those of the second. The sum and product of the roots of $x^2 + lx + m = 0$ are $-l$ and m respectively (Ch. II, § 16, or Ch. VI, § 1). Hence

$$(8) \qquad \begin{cases} x_1 + x_2 = -\frac{1}{2}(b - t), & x_1x_2 = \frac{1}{2}y_1 - (\frac{1}{2}by_1 - d)/t, \\ x_3 + x_4 = -\frac{1}{2}(b + t), & x_3x_4 = \frac{1}{2}y_1 + (\frac{1}{2}by_1 - d)/t. \end{cases}$$

Using also (5), we find at once that

$$(9) \qquad x_1 + x_2 + x_3 + x_4 = -b, \quad x_1x_2x_3x_4 = \tfrac{1}{4}y_1^2 - (\tfrac{1}{4}y_1^2 - e) = e,$$

$$(10) \quad x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4 = x_1x_2 + (x_1 + x_2)(x_3 + x_4) + x_3x_4 = c,$$

$$(11) \quad x_1x_2x_3 + x_1x_2x_4 + x_1x_3x_4 + x_2x_3x_4 = x_1x_2(x_3 + x_4) + x_3x_4(x_1 + x_2) = -d.$$

It follows from Ex. 3, p. 40 that (9)–(11) hold also when there is no root y_1 for which $t \neq 0$.

For any quartic equation (1), *the sum of the roots is* $-b$, *the sum of the products of the roots two at a time is* c, *the sum of the products three at a time is* $-d$, *the product of all four is* e.

A proof based upon more fundamental principles is given in Ch. VI, § 1.

3. Roots of the Resolvent Cubic Equation. These are

$$(12) \qquad y_1 = x_1x_2 + x_3x_4, \quad y_2 = x_1x_3 + x_2x_4, \quad y_3 = x_1x_4 + x_2x_3.$$

The first relation follows from (8). If, instead of y_1, another root of (6) be employed as in § 1, quadratic equations different from (7) are obtained, such however that their four roots are x_1, x_2, x_3, x_4, paired in a new way. This leads us to expect that y_2 and y_3 in (12) are the remaining roots of cubic (6). To give a formal proof, note that, by (9)–(11),

$$(13)\quad \begin{cases} y_1 + y_2 + y_3 = c, \\ y_1y_2+y_1y_3+y_2y_3 = (x_1+x_2+x_3+x_4)(x_1x_2x_3+\cdots+x_2x_3x_4) - 4\,x_1x_2x_3x_4 \\ \qquad = bd - 4\,c, \\ y_1y_2y_3 = (x_1x_2x_3 + \cdots)^2 + x_1x_2x_3x_4\{(x_1+\cdots)^2 - 4\,(x_1x_2+\cdots)\} \\ \qquad = d^2 + e\,(b^2 - 4\,c). \end{cases}$$

Hence by Ex. 7, p. 32, or by Ch. VI, §1, y_1, y_2, y_3 are the roots of (6).

EXERCISES

1. Why is it sufficient for the last proof to verify merely the first two relations (13)?

2. In Lagrange's solution of quartic (1), we begin by showing that the numbers (12) are the roots of cubic (6) by using (13) and the theorem of § 2. Let a root y_1 be found. Then we obtain $x_1x_2 = z_1$ and $x_3x_4 = z_2$ as the roots of $z^2 - y_1z + e = 0$. Next, $x_1 + x_2$ and $x_3 + x_4$ are found from

$$(x_1 + x_2) + (x_3 + x_4) = -b, \quad z_2(x_1 + x_2) + z_1(x_3 + x_4) = -d.$$

Hence x_1 and x_2, x_3 and x_4 are found by solving quadratic equations. Give the details of this work.

3. If the t corresponding to each root of (6) is zero, equation (1) has all its roots equal. For, by (3), the y's all equal $c - \frac{1}{4}b^2$. By (13), $3\,y_1 = c$, $3\,y_1^2 = bd - 4\,e$. Hence $c = \frac{3}{8}b^2$, $\frac{3}{64}b^4 = bd - 4\,e$. Eliminating e between the latter and $(\frac{1}{8}b^2)^3 = y_1^3 = b^2e - 4\,ce + d^2$, which follows from $y_1 = c - \frac{1}{4}b^2$ and (13), we get $(\frac{1}{16}b^3 - d)^2 = 0$. Then (1) equals $(x + \frac{1}{4}b)^4 = 0$.

4. Prove that Ex. 3 is true by showing that $t^2 = (x_1 + x_2 - x_3 - x_4)^2$.

5. Solve $x^3 + px + q = 0$ $(p \neq 0)$ by choosing c so that the quartic

$$(x - c)(x^3 + px + q) = 0$$

shall have as its resolvent cubic (6) one reducible to the form z^3 = constant. Here (6) is

$$y^3 - py^2 + c(cp + 3\,q)y - c^2p^2 - 2\,cpq - q^2 + c^3q = 0.$$

To remove the second term, set $y = z + p/3$. We get

$$z^3 + Az + c^3q - \tfrac{2}{3}c^2p^2 - cpq - q^2 - \tfrac{2}{27}p^3 = 0,$$

where $A = pc^2 + 3\,cq - \frac{1}{3}p^2$. We are to make $A = 0$; thus

$$\tfrac{1}{3}pc = -\tfrac{1}{2}q + \sqrt{R}, \qquad R = \frac{q^2}{4} + \frac{p^3}{27},$$

$$z^3 = -q(c^3 + cp + q) + 8\,R = \left(\frac{6}{p}\sqrt{R}\right)^3\left(-\tfrac{1}{2}q + \sqrt{R}\right),$$

since $c^3 + cp + q = 36\,cR/p^2$. Our quartic has the root c and hence by (8_1), with b replaced by $-c$, also the root $\frac{1}{2}(c + t) - c$, where $t^2 = c^2 - 4\,p + 4\,y$. Hence the given cubic has the root

$$\tfrac{1}{2}(t - c) = \sqrt{z - \tfrac{2}{3}p + \tfrac{1}{4}c^2} - \tfrac{1}{2}c,$$

which may be reduced to Cardan's form (*Amer. Math. Monthly*, 1898, p. 38).

4. Discriminants. Replacing y by $Y + c/3$ in (6), we get

$$Y^3 + PY + Q = 0, \tag{14}$$

in which

$$P = bd - 4\,e - \tfrac{1}{3}c^2, \quad Q = -b^2e + \tfrac{1}{3}bcd + \tfrac{8}{3}ce - d^2 - \tfrac{2}{27}c^3. \tag{15}$$

Hence (Ch. III, § 3),

$$(y_1 - y_2)^2(y_1 - y_3)^2(y_2 - y_3)^2 = -4\,P^3 - 27\,Q^2.$$

By (12)

$$\begin{aligned} y_1 - y_2 &= (x_1 - x_4)(x_2 - x_3), \\ y_1 - y_3 &= (x_1 - x_3)(x_2 - x_4), \\ y_2 - y_3 &= (x_1 - x_2)(x_3 - x_4). \end{aligned} \tag{16}$$

The discriminant Δ of the quartic (1) is defined to be

$$\Delta = (x_1 - x_2)^2(x_1 - x_3)^2(x_1 - x_4)^2(x_2 - x_3)^2(x_2 - x_4)^2(x_3 - x_4)^2. \tag{17}$$

It therefore equals the discriminant of (14):

$$\Delta = -4\,P^3 - 27\,Q^2. \tag{18}$$

Any quartic equation and its resolvent cubic have equal discriminants.

Some writers define the discriminant of (1) to be $\Delta/256$ and that of a cubic to be $-\Delta/27$. In suppressing these numerical factors, we have spared the reader a feat of memory, simplified the important relation between the discriminants of a quartic equation and its resolvent cubic, and moreover secured uniformity with most of the books to which we shall have occasion to refer the reader. Finally, we note that in applications to the theory of numbers, the insertion of the numerical factors is undesirable and in special cases unallowable (*cf. Bull. Amer. Math. Soc.*, vol. 13, 1906, p. 1).

EXERCISES

1. For $ax^4 + bx^3 + cx^2 + dx + e = 0$, $P = p/a^2$, $Q = q/a^3$, where

$$p = bd - 4\,ae - c^2/3, \quad q = -b^2e + \tfrac{1}{3}bcd + \tfrac{8}{3}ace - ad^2 - \tfrac{2}{27}c^3.$$

The discriminant is defined to be $a^6\Delta$; it equals $-4\,p^3 - 27\,q^2$.

2. If x and y are interchanged in

$$f = ax^4 + bx^3y + cx^2y^2 + dxy^3 + ey^4,$$

a function is obtained which may also be derived from f by merely interchanging a with e, and b with d. Show that the latter interchanges leave p, q and the discriminant unaltered.

3. Since the sum $Y_1 + Y_2 + Y_3$ of the roots of a reduced cubic is zero,

$$Y_1 = \tfrac{1}{3}(Y_1 - Y_2) + \tfrac{1}{3}(Y_1 - Y_3), \ldots,$$

and any root and hence any function of the roots is expressible as a function of the differences of the roots. Thus P and Q in (15) are functions of $Y_1 - Y_2$, etc., and hence of $y_1 - y_2$, etc. Using (16), show that p and q equal polynomials in the differences of $x_1, \ldots, x_4$.

4. When x is replaced by $x + ty$, let f of Ex. 2 become

$$f' = a'x^4 + b'x^3y + \cdots + e'y^4.$$

Show by Ex. 3 that p and q equal the corresponding functions

$$p' = b'd' - 4a'e' - c'^2/3, \quad q' = -b'^2e' + \cdots.$$

5. The results in Exs. 2 and 4 are special cases (used in a short proof) of a general theorem: When x is replaced by $lx + my$ and y by $rx + sy$, let f become f'. Then, using the notations of Ex. 4, we have $p' = D^4p$, $q' = D^6q$, where $D = ls - mr$. Hence p and q are called *invariants* of f. Verify the theorem for the case when x is replaced by lx, y by y.

6. The discriminant is an invariant and the factor is D^{12}.

7. Using $a_0x^4 + 4a_1x^3y + 6a_2x^2y^2 + 4a_3xy^3 + a_4y^4$ in place of the former f, show that $p = -4I$, $q = 16J$, where

$$I = a_0a_4 - 4a_1a_3 + 3a_2^2, \quad J = a_0a_2a_4 + 2a_1a_2a_3 - a_0a_3^2 - a_1^2a_4 - a_2^3.$$

In (14) set $Y = 2z/a$; then $z^3 - Iz + 2J = 0$. The discriminant is

$$256(I^3 - 27J^2).$$

5. Descartes' Solution of the Quartic Equation. Replacing x by $z - b/4$ in the general quartic (1), we obtain a *reduced* quartic equation

$$z^4 + qz^2 + rz + s = 0, \tag{19}$$

lacking the term with z^3. We shall prove that we can express the left member of (19) as the product of two quadratic factors*

$$(z^2 + 2kz + l)(z^2 - 2kz + m) = z^4 + (l + m - 4k^2)z^2 + 2k(m - l)z + lm.$$

* If the coefficients of z be denoted by k and $-k$ (as is usually done), the expressions (23) for the roots must be divided by 2. But the identification with Euler's solution is then not immediate.

The conditions are

$$l + m - 4k^2 = q, \quad 2k(m - l) = r, \quad lm = s.$$

If $k \neq 0$, the first two give

$$2m = q + 4k^2 + \frac{r}{2k}, \quad 2l = q + 4k^2 - \frac{r}{2k}.$$

Then $lm = s$ gives

$$64k^6 + 32qk^4 + 4(q^2 - 4s)k^2 - r^2 = 0. \tag{20}$$

The latter may be solved as a cubic equation for k^2. Any root $k^2 \neq 0$ gives a pair of quadratic factors of (19):

$$z^2 \pm 2kz + \tfrac{1}{2}q + 2k^2 \mp \frac{r}{4k}. \tag{21}$$

The 4 roots of these two quadratic functions are the 4 roots of (19). If $q = r = s = 0$, every root of (20) is zero and the discussion is not valid; but the quadratic factors are then evidently z^2, z^2.

EXERCISES

1. For $z^4 - 3z^2 + 6z - 2 = 0$, (20) becomes

$$64k^6 - 3 \cdot 32k^4 + 4 \cdot 17k^2 - 36 = 0.$$

The value $k^2 = 1$ gives the factors $z^2 + 2z - 1$, $z^2 - 2z + 2$, with the roots $-1 \pm \sqrt{2}$, $1 \pm \sqrt{-1}$.

2. Solve $z^4 - 2z^2 - 8z - 3 = 0$.
3. Solve $z^4 - 10z^2 - 20z - 16 = 0$.
4. Solve $x^4 - 8x^3 + 9x^2 + 8x - 10 = 0$.

6. Symmetrical Form of Descartes' Solution. To obtain this symmetrical form, we use all three roots k_1^2, k_2^2, k_3^2 of (20). Then

$$k_1^2 + k_2^2 + k_3^2 = -\tfrac{1}{2}q, \quad k_1^2k_2^2k_3^2 = r^2/64.$$

It is at our choice as to which square root of k_1^2 is denoted by $+k_1$ and which by $-k_1$, and likewise as to $\pm k_2$, $\pm k_3$. For our purposes any choice of these signs is suitable provided the choice give

$$k_1k_2k_3 = -r/8. \tag{22}$$

Let $k_1 \neq 0$. The quadratic function (21) is zero for $k = k_1$ if

$$(z \pm k_1)^2 = -\frac{q}{2} - k_1^2 \pm \frac{r}{4k_1} = k_2^2 + k_3^2 \mp \frac{8k_1k_2k_3}{4k_1} = (k_2 \mp k_3)^2.$$

Hence the four roots of the quartic equation (19) are

$$(23)\quad k_1 + k_2 + k_3, \quad k_1 - k_2 - k_3, \quad -k_1 + k_2 - k_3, \quad -k_1 - k_2 + k_3.$$

Writing $k^2 = y$, we see that, if y_1, y_2, y_3 are the roots of

$$(24)\qquad 64\, y^3 + 32\, qy^2 + 4\,(q^2 - 4\, s)y - r^2 = 0,$$

then the roots of (19) are the four values

$$(25)\qquad z = \sqrt{y_1} + \sqrt{y_2} + \sqrt{y_3},$$

obtained by using all of the combinations of the square roots for which, by (22),

$$(26)\qquad \sqrt{y_1}\,\sqrt{y_2}\,\sqrt{y_3} = -r/8.$$

We have deduced Euler's solution (Ex. 1) from Descartes'.

EXERCISES

1. Assume with Euler that quartic (19) has a root of the form (25). Square (25), transpose the terms free of radicals, square again, and show that

$$z^4 - 2\,(y_1 + y_2 + y_3)\, z^2 - 8\, z\, \sqrt{y_1}\,\sqrt{y_2}\,\sqrt{y_3} + (y_1 + y_2 + y_3)^2$$
$$- 4\,(y_1y_2 + y_1y_3 + y_2y_3) = 0.$$

From the relations obtained by identifying this with (19), show that y_1, y_2, y_3 are the roots of the cubic (24) and that (26) holds.

2. Solve Exs. 1–4 of the preceding set by use of (23).

3. In the theory of inflexion points of a plane cubic curve occurs the quartic equation $z^4 - Sz^2 - \frac{4}{3}\, Tz - \frac{1}{12}\, S^2 = 0$. Show that (24) now becomes

$$\left(y - \frac{S}{6}\right)^3 = C, \quad C \equiv \left(\frac{T}{6}\right)^2 - \left(\frac{S}{6}\right)^3,$$

and that the roots of the quartic are

$$\pm\sqrt{\tfrac{1}{6}\, S + \sqrt[3]{C}} \pm \sqrt{\tfrac{1}{6}\, S + \omega\, \sqrt[3]{C}} \pm \sqrt{\tfrac{1}{6}\, S + \omega^2\, \sqrt[3]{C}},$$

where the signs are to be chosen so that the product of the three summands equals $+\,T/6$. Here ω is an imaginary cube root of unity.

4. The discriminant Δ of the quartic equation (19) equals the quotient of the discriminant D of (24) by 4^6. For, the six differences of the roots (23) are $2\,(k_1 \pm k_2)$, $2\,(k_1 \pm k_3)$, $2\,(k_2 \pm k_3)$. Thus $\Delta = 4^6\, L$, where

$$L = (k_1^2 - k_2^2)^2(k_1^2 - k_3^2)^2(k_2^2 - k_3^2)^2 = (y_1 - y_2)^2(y_1 - y_3)^2(y_2 - y_3)^2.$$

By definition, $D = 64^4 L$. Hence $D = 4^6\, \Delta$.

5. Give a second proof of Ex. 4 by setting $y = z/4$ in (24) and then $z = Y - 2q/3$. We obtain (14), in which now $b = 0$, $c = q$, $d = r$, $e = s$. The discriminant of (14) equals Δ. Hence $\Delta = (z_1 - z_2)^2 \cdots = 4^6 L = D/4^6$.

6. If a quartic equation has two pairs of conjugate imaginary roots, its discriminant Δ is positive. Hence, if $\Delta < 0$, there are exactly two real roots.

7. Theorem.* *A quartic equation* (19) *with* q, r, s, *real*, $r \neq 0$, *and with the discriminant* Δ, *has*

4 distinct real roots if q and $4s - q^2$ are negative and $\Delta > 0$,
no real root if q and $4s - q^2$ are not both negative and $\Delta > 0$,
2 distinct real and 2 imaginary roots if $\Delta < 0$,
at least 2 equal real roots if $\Delta = 0$.

Since the constant term of the cubic equation (24) is negative, at least one of its roots is a positive real number. Let, therefore, $y_1 > 0$, so that $y_2 y_3 > 0$. Thus $k_1 = \sqrt{y_1}$ is real. There are four possible cases to consider.

(*a*) y_2 and y_3 positive. Then each $k_j = \sqrt{y_j}$ is real and the roots (23) of the quartic equation are all real.

(*b*) $y_2 = y_3 < 0$. Then $k_2 = \pm k_3$ is a pure imaginary. If $k_2 = k_3$, the first two roots (23) are imaginary and the last two are real and equal. If $k_2 = -k_3$, the reverse is true.

(*c*) y_2 and y_3 distinct and negative. The roots (23) are all imaginary.

(*d*) y_2 and y_3 conjugate imaginaries. Then k_2 is imaginary and conjugate with either k_3 or $-k_3$, so that one of the numbers $k_2 + k_3$ and $k_2 - k_3$ is real and the other imaginary. Just two of the roots (23) are real.

Now, if $\Delta = 0$, at least two y's are equal by Ex. 4 of the last set. Thus we have case (*b*) or a special case of (*a*). In either case, the quartic has at least two equal roots, by (17), and they are real in both cases.

Henceforth, let $\Delta \neq 0$. By the same Ex. 4, Δ has the same sign as the discriminant D of the cubic equation (24). If $\Delta < 0$, we have case (*d*). Finally, let $\Delta > 0$, so that y_1, y_2, y_3 are real. If q is negative and $q^2 - 4s$ is positive, equation (24) has alternately positive and negative coefficients and hence has no negative root, so that we have case (*a*). But if q and $4s - q^2$ are not both negative, the coefficients are not alternately positive and negative, so that the roots y_1, y_2, y_3 are not all positive,** and we have case (*c*).

* Proved by Lagrange by use of the equation whose six roots are the squares of the differences of the roots of (19), *Résolution des équations numériques*, 3d ed., p. 42.

** The coefficients are $-(y_1 + y_2 + y_3)$, $y_1 y_2 + y_1 y_3 + y_2 y_3$, $-y_1 y_2 y_3$.

EXERCISES

1. Apply this theorem to the quartic equations in Exs. 1–4, p. 43.
2. Verify that a quartic equation (19) with two pairs of equal imaginary roots has $r = 0$. Deduce the last case of the theorem.
3. Why does the theorem imply its converse?

† CHAPTER V

The Fundamental Theorem of Algebra

1.† Theorem. *Every equation with complex coefficients*

$$f(z) \equiv z^n + a_1 z^{n-1} + \cdots + a_n = 0 \tag{1}$$

has a complex (real or imaginary) root.

For $n = 2$, 3, or 4, we have proved this theorem by actually solving the equation. But for $n \geqq 5$, the equation cannot in general be solved algebraically, *i.e.*, in terms of radicals.

We shall first treat the case in which all of the coefficients are real. Relying upon geometrical intuition, we have seen in Exs. 3, 5, p. 14, that there is a real root if n is odd, or if both n is even and a_n is negative. But, as in the cases of certain quadratic equations and $z^4 + z^2 + 5 = 0$, an equation of even degree may have no real root. No proof of the theorem for all cases has been made by such elementary methods.

The proof here given of the theorem that any equation with real coefficients has a complex root is essentially the first proof by Gauss (1799 and simplified by him in 1849).

We are to prove that there exists a complex number $z = x + yi$ such that $f(z) = 0$. We may write

$$f(z) = X + Yi, \tag{2}$$

where X and Y are polynomials in x and y with real coefficients. We are to show that there exist real numbers x and y such that

$$X = 0, \quad Y = 0. \tag{3}$$

For example, if $f(z) = z^4 - 4z^3 + 9z^2 - 16z + 20$, then

$$X = x^4 - 6x^2y^2 + y^4 - 4x^3 + 12xy^2 + 9x^2 - 9y^2 - 16x + 20,$$
$$\tfrac{1}{2}Y = 2x^3y - 2xy^3 - 6x^2y + 2y^3 + 9xy - 8y.$$

The graph of $Y = 0$ is the x-axis ($y = 0$) and the graph (indicated by the dotted curve in Fig. 16, asymptotic to the lines $x = 1$ and $y = \pm x$) of

$$2(x - 1)y^2 = 2x^3 - 6x^2 + 9x - 8.$$

Note that there is no real y for x between 1 and 1.73. Since $X = 0$ is a quadratic equation in y^2, its graph is readily drawn. There is no real y for $x = 0.05$ and 1.6

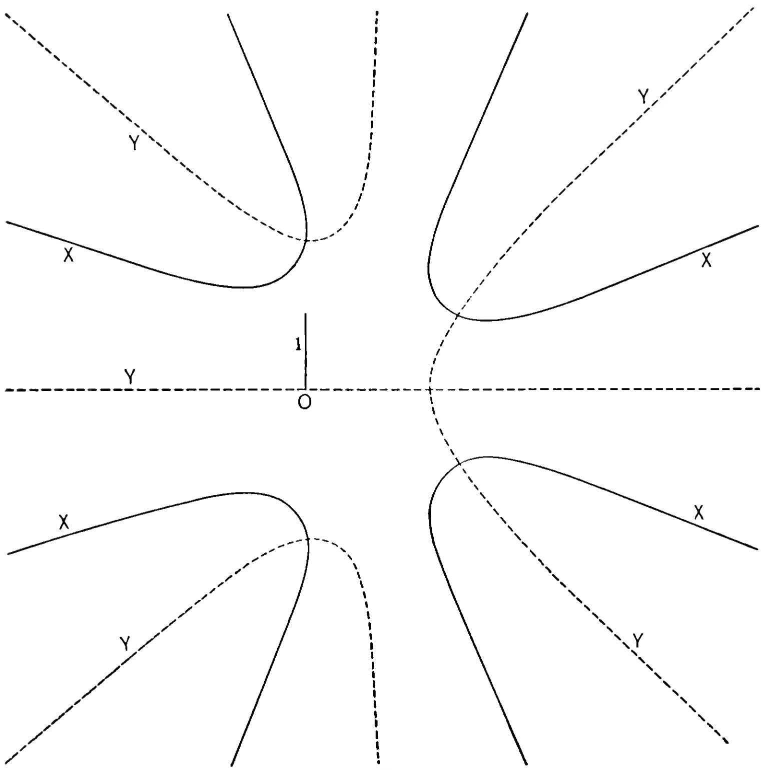


Fig. 16

and the intermediate values. Cases in which the values of y^2 are positive and rational are

x	-4	-2	-1	0	2	3
y^2	5, 148	2.5, 54.5	2, 25	4, 5	1, 8	1, 26

The graphs cross at the points (0, 2), (0, -2), (2, 1), (2, -1), and the roots of $f(z) = 0$ are $z = \pm 2i, 2 \pm i$.

We shall employ also the trigonometric form of z:

$$z = r(\cos\theta + i\sin\theta), \tag{4}$$

where $0 \leqq \theta < 2\pi$. Set $t = \tan \frac{1}{2}\theta$. Then

$$\frac{2t}{1+t^2} = \frac{2\tan\frac{1}{2}\theta}{\sec^2\frac{1}{2}\theta} = 2\sin\tfrac{1}{2}\theta \cdot \cos\tfrac{1}{2}\theta = \sin\theta,$$

$$\tan\theta = \frac{2t}{1-t^2}, \quad \cos\theta = \frac{\sin\theta}{\tan\theta} = \frac{1-t^2}{1+t^2}.$$

Thus

$$z = \frac{r(1+ti)^2}{1+t^2}.$$

Hence by (1) and (2),

$$(1+t^2)^n(X+Yi) = r^n(1+ti)^{2n} + a_1r^{n-1}(1+ti)^{2n-2}(1+t^2) + \cdots + a_n(1+t^2)^n.$$

Expanding the terms on the right by the binomial theorem, we get

$$(5) \qquad X = \frac{F(t)}{(1+t^2)^n}, \quad Y = \frac{G(t)}{(1+t^2)^n},$$

where $F(t)$ is a polynomial in t of degree $2n$, and $G(t)$ a polynomial in t of degree less than $2n$, each with coefficients involving r integrally.

Each point (x, y), representing (Ch. II, § 8) a complex number $z = x + yi$ having the modulus r, lies on the circle $x^2 + y^2 = r^2$ with radius r and center at the origin of the rectangular coördinate system. To find the points on this circle for which $X = 0$ or $Y = 0$, we solve $F(t) = 0$ or $G(t) = 0$ (in which r is now a constant), and note that to each real root t corresponds a single real value of $\sin\theta$ and a single real value of $\cos\theta$, consistent with that of $\sin\theta$, and hence a single point ($x = r\cos\theta$, $y = r\sin\theta$). But an equation of degree $2n$ has at most $2n$ distinct roots (Ch. I, § 15). Since the degree of $G(t)$ is less than that of the denominator of Y in (5), the root $t = \infty$ of $Y = 0$ must be considered in addition to the roots of $G(t) = 0$ already examined; for $t = \infty$, $\theta = \pi$ and the point is $(-r, 0)$. Thus neither X nor Y is zero for more than $2n$ points of the circle with center at the origin and a given radius r. By proper choice of r, this circle will have an arc lying within any given region of the plane. Hence *neither X nor Y is zero at all points of a region of the plane.*

From (4) and De Moivre's Theorem (Ch. II, § 10), we have

$$z^k = r^k(\cos k\theta + i\sin k\theta).$$

Hence, by (1) and (2),

$$Y = r^n \sin n\theta + a_1r^{n-1}\sin(n-1)\theta + a_2r^{n-2}\sin(n-2)\theta + \cdots + a_{n-1}r\sin\theta.$$

Let g be the greatest of the numerical values of $a_1, \ldots, a_{n-1}$. Then, if $|D|$ denotes the numerical value of the real number D,

$$Y = r^n(\sin n\theta + D), \quad |D| \leqq g\left(\frac{1}{r} + \frac{1}{r^2} + \cdots + \frac{1}{r^{n-1}}\right) < g\left(\frac{\frac{1}{r}}{1 - \frac{1}{r}}\right),$$

provided $r > 1$. If c is a positive constant < 1 and if $r > 1 + g/c$, then $|D| < c$. Hence for all angles θ for which $\sin n\theta$ is numerically greater than c, Y has the same sign as its first term $r^n \sin n\theta$ when r exceeds the constant $1 + g/c$.

In our example, we have

$$Y = r^4 \sin 4\,\theta - 4\,r^3 \sin 3\,\theta + 9\,r^2 \sin 2\,\theta - 16\,r \sin \theta.$$

The limit $1 + 16/c$ for r exceeds 17 and is larger than is convenient for a drawing. But for $r \geqq 10$,

$$Y = r^4(\sin 4\,\theta + D), \quad |D| \leqq \frac{4}{r} + \frac{9}{r^2} + \frac{16}{r^3} \leqq 0.4 + 0.09 + 0.016.$$

Taking $c = 0.506 = \sin 30°\ 24'$, let C be the number of radians in $7°\ 36'$. Thus $c = \sin 4\,C$. The positive angles θ ($\theta < 2\,\pi$) for which $\sin 4\,\theta$ exceeds $\sin 4\,C$ numerically are those between C and $\frac{1}{4}\pi - C$, between $\frac{1}{4}\pi + C$ and $\frac{1}{2}\pi - C$, between $\frac{1}{2}\pi + C$ and $\frac{3}{4}\pi - C, \ldots,$ between $\frac{7}{4}\pi + C$ and $2\,\pi - C$. For any such angle θ and for $r \geqq 10$, Y has the same sign as $\sin 4\,\theta$ and hence is alternately positive and negative in these successive intervals, the solid arcs in Fig. 17. Denote by $0, 1, 2, \ldots, 7$ the points on the circle with center at the origin and radius 10 whose angles θ are $0, \frac{\pi}{4}, \frac{2\,\pi}{4}, \ldots, \frac{7\,\pi}{4}$, respectively.

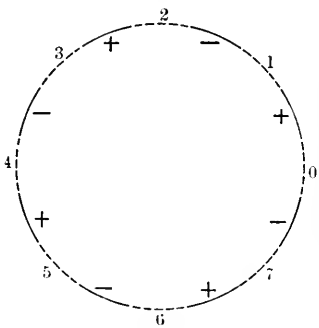


Fig. 17

In the general case, denote by $0, 1, 2, \ldots, 2\,n - 1$ the points with the angles

$$0, \frac{\pi}{n}, \frac{2\,\pi}{n}, \ldots, \frac{(2\,n-1)\,\pi}{n}$$

on the circle with center at the origin and radius a constant r exceeding the above value $1 + g/c$. Let nC be the positive angle $< \pi/2$ for which $\sin nC = c$. We define the neighborhood of our kth point of division on

the circle to be the arc bounded by the points whose angles are $k\pi/n - C$ and $k\pi/n + C$. In Fig. 17 for our example with $n = 4$, each neighborhood is indicated by a dotted arc. In the successive arcs (marked by solid arcs) between the neighborhoods, Y is alternately positive and negative, since it has in each the same sign as $\sin n\theta$.

It is easily seen that $\sin \theta, \sin 2\,\theta, \ldots, \sin n\theta$ are continuous functions of θ (a fact presupposed in interpolating between values read from a table of sines). Since r is now a constant, Y is therefore a continuous function of θ, and has a single value for each value of θ. But Y has opposite signs at the two ends of the neighborhood of any one of our points of division on the circle. Hence (as in Ch. I, § 12), Y is zero for some point within each neighborhood, and at just one such point, since Y was shown to vanish at not more than $2\,n$ points of a circle with center at the origin. We shall denote the points on the circle at which Y is zero by

$$P_0,\ P_1,\ \ldots,\ P_{2n-1}.$$

For our example, these points $P_0, \ldots, P_7$ are given in Fig. 18, which shows more of the graph of $Y = 0$ than was given in Fig. 16, but now shows it with the

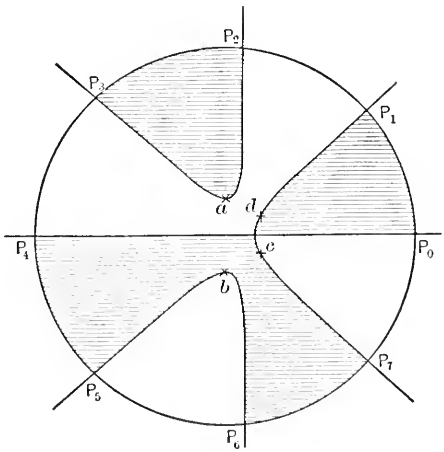


Fig. 18

scale of length reduced in the ratio 4 to 1 (to have a convenient circle of radius 10). We have shaded the regions in which, as next proved, Y is positive.

Let the constant r be chosen so large that X also has the same sign as its first term $r^n \cos n\theta$, for θ not too near one of the values $\pi/(2n)$, $3\pi/(2n)$, $5\pi/(2n)$, . . . , for which $\cos n\theta = 0$. Since these values correspond to the middle points of the arcs $(\widehat{01})$, $(\widehat{12})$, . . . , no one of them lies in a neighborhood of a division point 0, 1, ... Now $\cos n\theta = +1$ or -1 when θ is an even or an odd multiple of π/n, respectively. Hence X is positive in the neighborhood of the division points $0, 2, 4, \ldots, 2n-2$ and thus at $P_0, P_2, P_4, \ldots$, but negative in that of $1, 3, 5, \ldots, 2n-1$ and thus at $P_1, P_3, P_5, \ldots$.

We saw that Y is not zero throughout a region of the plane. Hence there is a region in which Y is everywhere positive (called a positive region), and perhaps regions in which Y is everywhere negative (called negative regions), while Y is zero on the boundary lines.

In Fig. 18 for our example, there are three positive (shaded) regions, the two with a single point in common being considered distinct, and three negative (unshaded) regions. Consider that part of the boundary of P_2P_3a which lies inside the circle. At every point of it, Y is zero. Now X is negative at P_3 and positive at P_2 and hence is zero at some intermediate point a on this boundary. Hence at a both X and Y are zero, so that a represents a complex root (in fact, $2i$) of $f(z) = 0$.

To extend the last argument to the general case, let R be the part inside our circle of a positive region having the points P_{2h} and P_{2h+1} on its boundary. The points of arc $P_{2h}P_{2h+1}$ may be the only boundary points of R lying on the circle (as for P_2P_3a and P_0P_1d in Fig. 18), or else its boundary includes at least another such arc $P_{2k}P_{2k+1}$ (as shaded region $P_4P_5bP_6P_7c$ in Fig. 18). In the first case, X and Y are both zero at some point (a or d) on the inner boundary, since X is negative at P_{2h+1} and positive at P_{2h} and hence zero at an intermediate point. In the second case, a point moving from P_{2h} to P_{2h+1} along the smaller included arc and then along the inner boundary of R until it first returns to the circle arrives at a point P_{2k} of even subscript (as in the case of $P_4P_5bP_6$). Indeed, if a person travels as did the point, he will always have the region R at his left and hence will pass from P_{2k} to P_{2k+1} and not *vice versa*. Since X is negative at P_{2h+1} and positive at P_{2k}, it (as also Y) is zero at some point b on the part of the inner boundary of R joining these two points. Hence b represents a root of $f(z) = 0$. Thus in either of the two possible cases, the equation has a root, real or imaginary.

2.† It remains to prove that an equation $F(z) = 0$, not all of whose coefficients are real, has a complex root. By separating each imaginary coefficient into its real and purely imaginary parts, we have $F(z) = P + Qi$, where P and Q are polynomials in z with real coefficients. Let $G(z) = P - Qi$. The equation

$$F(z) \cdot G(z) \equiv P^2 + Q^2 = 0$$

has real coefficients and hence has a complex root $z = a + bi$. If this is a root of $F(z) = 0$, our theorem is proved. If it is not, then $G(a + bi) = 0$. Then by Ch. II, § 16, $F(a - bi) = 0$, and the given equation has the root $a - bi$.

EXERCISES

1.† For $z^3 = 11 + 2i$, draw the graphs of $X = 0$, $Y = 0$ and locate the three roots of the cubic equation in z.

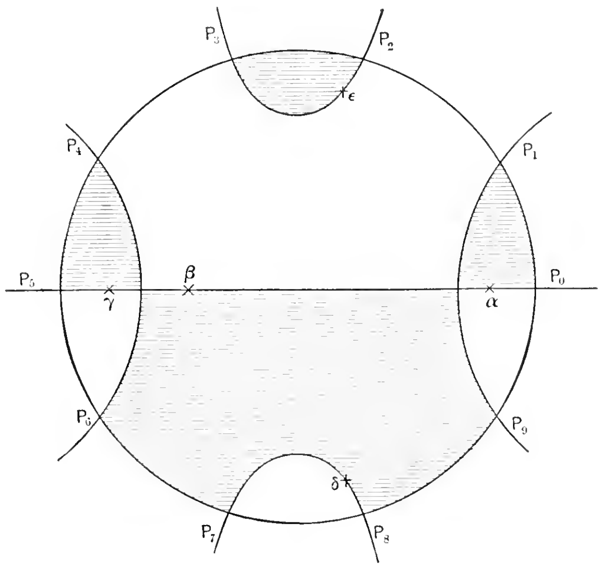


Fig. 19

2.† For $z^5 - 4z - 2 = 0$, $Y = r^5 \sin 5\theta - 4r \sin \theta$. Using polar coördinates, show that the graph of $Y = 0$ gives the boundaries of the regions in Fig. 19: first plot the horizontal line corresponding to $\sin \theta = 0$, and then, using various angles θ $(\theta \neq 0, \pi)$, find by logarithms the corresponding positive r from

$$r^4 = \frac{4 \sin \theta}{\sin 5\theta}.$$

To find the points on these boundaries ($Y = 0$) for which also

$$X \equiv r^5 \cos 5\,\theta - 4\,r \cos \theta - 2 = 0,$$

replace r^4 by the earlier expression. We get

$$4\,r(\sin \theta \cos 5\,\theta - \cos \theta \sin 5\,\theta) = 2 \sin 5\,\theta, \quad r = -\frac{\sin 5\,\theta}{2 \sin 4\,\theta}.$$

Comparing the fourth power of this fraction with that for r^4, we get

$$\sin^5 5\,\theta = 64 \sin \theta \sin^4 4\,\theta,$$

which holds for $\theta = 85° \; 21' \; 30''$ or its negative. We then get r and therefore the roots

$$\epsilon, \delta = 0.11679 \pm 1.4385\,i.$$

On the horizontal line are three real roots, best found by methods of approximation given later:

$$\alpha = 1.518512, \quad \beta = -0.5084994, \quad \gamma = -1.2435964.$$

(H. Weber and J. Wellstein, *Encyklopädie der Elementar-Mathematik*, ed. 1, I, p. 212, p. 296.)

3.† Other References. For proofs of the fundamental theorem by Gauss, Cauchy and Gordan, see Netto, *Vorlesungen über Algebra*, I, p. 25, p. 173. The shortest proofs are by the use of the theory of functions of a complex variable, and may be found in texts on that subject. For an algebraic proof resting upon the theory of functions of a real variable, see Weber, *Lehrbuch der Algebra*, 2d ed., vol. 1, pp. 119–142. See also *Monographs on Topics of Modern Mathematics*, 1911, p. 201, edited by Young (article by Huntington). In the *Amer. Math. Monthly*, vol. 10 (1903), p. 159, Moritz has pointed out hidden assumptions in various incomplete proofs.

CHAPTER VI

Elementary Theorems on the Roots of an Equation

1. Relations between the Roots and the Coefficients. Given an equation in x of degree n, we can divide its members by the coefficient of x^n and obtain an equation of the form

$$(1) \qquad f(x) \equiv x^n + p_1 x^{n-1} + p_2 x^{n-2} + \cdots + p_n = 0.$$

By the fundamental theorem of algebra (Ch. V), it has a root α_1, and its quotient by $x - \alpha_1$ has a root α_2, etc. Thus

$$(2) \qquad f(x) \equiv (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n),$$

identically in x. Since the polynomial has n linear factors, each having one root, we shall say that the equation has n roots. These may not all be distinct; exactly m of them equal α_1, if α_1 is a root of multiplicity m, *i.e.*, if exactly m of the linear factors in (2) equal $x - \alpha_1$. Next,

$$(x - \alpha_1)(x - \alpha_2) \equiv x^2 - (\alpha_1 + \alpha_2)x + \alpha_1\alpha_2,$$

$$(x-\alpha_1)(x-\alpha_2)(x-\alpha_3) \equiv x^3 - (\alpha_1+\alpha_2+\alpha_3)x^2 + (\alpha_1\alpha_2+\alpha_1\alpha_3+\alpha_2\alpha_3)x - \alpha_1\alpha_2\alpha_3.$$

Thus for $n = 2$ or 3, we see that the product (2) equals

$$(3) \quad x^n - (\alpha_1 + \cdots + \alpha_n)x^{n-1} + (\alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3 + \cdots + \alpha_{n-1}\alpha_n)x^{n-2} \\ - (\alpha_1\alpha_2\alpha_3 + \alpha_1\alpha_2\alpha_4 + \cdots + \alpha_{n-2}\alpha_{n-1}\alpha_n)x^{n-3} + \cdots + (-1)^n\alpha_1\alpha_2\cdots\alpha_n.$$

Multiplying this by $x - \alpha_{n+1}$, we readily verify that the product is a function which may be derived from (3) by changing n into $n + 1$. It therefore follows by mathematical induction that (2) and (3) are identical. Hence (1) and (3) are identical, so that

$$(4) \qquad \begin{aligned} \alpha_1 + \alpha_2 + \cdots + \alpha_n &= -p_1, \\ \alpha_1\alpha_2 + \alpha_1\alpha_3 + \cdots + \alpha_{n-1}\alpha_n &= p_2, \\ \alpha_1\alpha_2\alpha_3 + \alpha_1\alpha_2\alpha_4 + \cdots + \alpha_{n-2}\alpha_{n-1}\alpha_n &= -p_3, \\ \cdots\cdots\cdots\cdots\cdots\cdots& \\ \alpha_1\alpha_2 \cdots \alpha_{n-1}\alpha_n &= (-1)^n p_n. \end{aligned}$$

For $n = 3$ and $n = 4$, the complete formulæ were given and proved otherwise in Ex. 7, p. 32 and Ch. IV, § 2.

In an equation in x of degree n, in which the coefficient of x^n is unity, the sum of the roots equals the negative of the coefficient of x^{n-1}, the sum of the products of the roots two at a time equals the coefficient of x^{n-2}, the sum of the products of the roots three at a time equals the negative of the coefficient of x^{n-3}, etc.; finally, the product of the roots equals the constant term or its negative according as n is even or odd.

For example, in a cubic equation having the roots 2, 2, 5, the coefficient of x equals $2 \cdot 2 + 2 \cdot 5 + 2 \cdot 5 = 24$.

Given an equation $a_0x^n + a_1x^{n-1} + \cdots = 0$, we first divide by a_0 and then apply the theorem to the resulting equation. Thus the sum of the roots equals $-a_1/a_0$.

EXERCISES

1. Find the quartic equation having 2 and -2 as double roots.
2. Find the remaining root in Exs. 1, 3, p. 9.
3. If a real cubic equation $x^3 - 6x^2 + \cdots = 0$ has the root $1 + \sqrt{-5}$, what are the remaining roots?
4. Form by the theorem the equations in Exs. 3, 4, p. 15.
5. Given that $x^4 - 2x^3 - 5x^2 - 6x + 2 = 0$ has the root $2 - \sqrt{3}$, find another root and, by using the sum and product of the four roots, form the quadratic equation for the remaining two roots (avoid division).
6. Find, by use of (4), the roots of $x^4 - 6x^3 + 13x^2 - 12x + 4 = 0$, given that it has two double roots.
7. Solve $x^3 - 3x^2 - 13x + 15 = 0$, with roots in arithmetical progression.
8. Solve $4x^3 - 16x^2 - 9x + 36 = 0$, one root being the negative of another.
9. Solve $x^3 - 9x^2 + 23x - 15 = 0$, one root being triple another.
10. Solve $x^3 - 14x^2 - 84x + 216 = 0$, with roots in geometrical progression.
11. Solve $x^4 - 2x^3 - 21x^2 + 22x + 40 = 0$, with roots in arithmetical progression. Denote them by $c - 3b$, $c - b$, $c + b$, $c + 3b$.
12. Solve $x^4 - 6x^3 + 12x^2 - 10x + 3 = 0$, with a triple root.
13. Find a necessary and sufficient condition that

$$f(x) \equiv x^3 + p_1x^2 + p_2x + p_3 = 0$$

shall have one root the negative of another. Note that

$$(\alpha_2 + \alpha_3)(\alpha_1 + \alpha_3)(\alpha_1 + \alpha_2)$$

is obtained by substituting $x = -p_1$ in (2).

14. If for $n = 4$ the roots of (1) satisfy the relation $\alpha_1\alpha_2 = \alpha_3\alpha_4$, then $p_1^2p_4 = p_3^2$. Note that (4) gives

$$-p_3 = \alpha_1\alpha_2(\alpha_3 + \alpha_4) + \alpha_3\alpha_4(\alpha_1 + \alpha_2) = -p_1\alpha_1\alpha_2.$$

15. What is the coefficient of y^{n-1} in the equation $y^n + \cdots = 0$ whose roots are $\alpha_1 - h, \cdots, \alpha_n - h$, when the α's are the roots of (1)? For what value of

h is this coefficient zero? Hence to remove the second term of an equation by replacing x by $y + h$, what value of h must we take? Check by the binomial theorem.

16. Find the equation whose roots are the roots of $x^3 - 6x^2 + 4 = 0$ each diminished by 3. Remove the second term by transformation.

17. Prove the binomial theorem by taking the α's all equal in (2) and (3) and counting the number of terms in each coefficient of (3).

18. Using (1) and (2), show that

$$(1 - \alpha_1^2)(1 - \alpha_2^2) \cdots (1 - \alpha_n^2) = (1 + p_2 + p_4 + \cdots)^2 - (p_1 + p_3 + p_5 + \cdots)^2,$$

$$(1 + \alpha_1^2)(1 + \alpha_2^2) \cdots (1 + \alpha_n^2) = (1 - p_2 + p_4 - \cdots)^2 + (p_1 - p_3 + p_5 - \cdots)^2.$$

19. Since $x_1, \ldots, x_4$, determined by relations (8) of Ch. IV, give the correct values of the sums (9)–(11), they are the roots of the quartic equation. Why does this give a new solution of the quartic?

20. Using Ex. 6, p. 32, make a similar argument for the cubic.

2. Upper Limit to the Positive Roots.

For an equation

$$f(x) \equiv a_0x^n + a_1x^{n-1} + \cdots + a_n = 0 \qquad (a_0 \neq 0)$$

with real coefficients, we shall prove the

Theorem. *If $a_0, a_1, \ldots, a_{k-1}$ are each $\geqq 0$, while $a_k < 0$, and if G is the greatest of the numerical values of the negative coefficients, each real root is less than $1 + \sqrt[k]{G/a_0}$.*

For positive values of x, $f(x)$ is numerically greater than or equal to

$$a_0x^n - G(x^{n-k} + x^{n-k-1} + \cdots + x + 1)$$

$$= a_0x^n - G\left(\frac{x^{n-k+1} - 1}{x - 1}\right) = \frac{x^{n-k+1}\{a_0(x^k - x^{k-1}) - G\} + G}{x - 1}.$$

But, if $x > 1$, $x^k - x^{k-1} \geqq (x - 1)^k$. Hence if $x \geqq 1 + \sqrt[k]{G/a_0}$,

$$a_0(x^k - x^{k-1}) \geqq G, \quad f(x) \neq 0.$$

3. Another Upper Limit to the Roots.

If the numerical value of each negative coefficient be divided by the sum of all of the positive coefficients which precede it, the greatest quotient so obtained when increased by unity gives an upper limit to the positive roots of the equation.

If the coefficient of x^m is positive, we replace x^m by

$$(x - 1)(x^{m-1} + x^{m-2} + \cdots + x + 1) + 1.$$

The argument will be clearer if applied to a particular case:

$$f(x) = p_0x^5 - p_1x^4 + p_2x^3 + p_3x^2 - p_4x + p_5 = 0,$$

where each p_i is positive. Then $f(x)$ is the sum of the terms

$$\begin{array}{rrrrrr}
p_0(x-1)x^4 & + p_0(x-1)x^3 & + p_0(x-1)x^2 & + p_0(x-1)x & + p_0(x-1) & + p_0 \\
- p_1x^4 & & p_2(x-1)x^2 & + p_2(x-1)x & + p_2(x-1) & + p_2 \\
& & & p_3(x-1)x & + p_3(x-1) & + p_3 \\
& & & - p_4x & & p_5.
\end{array}$$

The sum of the terms in each column will be positive, if $x > 1$ and

$$p_0(x-1) - p_1 > 0, \quad (p_0 + p_2 + p_3)(x-1) - p_4 > 0,$$

since only in the first and fourth columns is there a negative part. These inequalities both hold if

$$x > 1 + \frac{p_1}{p_0}, \qquad x > 1 + \frac{p_4}{p_0 + p_2 + p_3}.$$

EXERCISES

Apply the methods of both § 2 and § 3 to find an upper limit u to the roots of

1. $4x^5 - 8x^4 + 22x^3 + 98x^2 - 73x + 5 = 0$. By § 2, $u = 1 + 73/4$. By § 3, $u = 3$, since $1 + 8/4 = 3$, $1 + 73/124 < 3$.
2. $x^5 + 4x^4 - 7x^2 - 40x + 1 = 0$. By § 2, $u = 1 + \sqrt[3]{40} = 4.42$. By § 3, $u = 9$.
3. $x^4 - 5x^3 + 7x^2 - 8x + 1 = 0$.
4. $x^7 + 3x^6 - 4x^5 + 5x^4 - 6x^3 - 7x^2 - 8 = 0$.
5. $x^7 + 2x^5 + 4x^4 - 8x^2 - 32 = 0$.
6. If A is the greatest of the numerical values of $a_1, \ldots, a_n$, each root is less than $1 + A/a_0$. In the proof in § 2, set $k = 1$ and replace G by A.
7. A lower limit to the negative roots of $f(x) = 0$ may be found by applying the above theorems to $f(-x) = 0$. To obtain a lower limit to the positive roots consider $f(1/x) = 0$.
8. Find a lower limit to the negative roots in Exs. 3, 4.
9. Find a lower limit to the positive roots in Ex. 5.

4. The Term "Divisor." In certain texts it is stated that the relation $\alpha_1 \alpha_2 \ldots \alpha_n = \pm p_n$ in (4) implies that "every root of an equation is a divisor of the absolute term." This statement is either trivial or else is not always true. It is trivial if it means merely that the absolute term can be divided by any root (that root being a complex number), yielding a quotient which is a complex number. For, in this sense division is always possible (except when the divisor is zero), and a root not zero is a divisor of any number whatever. The statement quoted was certainly not meant in this trivial sense, with no special force. The only other sense, familiar to the reader, in which a constant is said to be a divisor

of another constant is the following: An *integer* r is a divisor of an *integer* p if p/r is an *integer*, so that $p = rq$, where q is an integer. For example, 4 is a divisor of 12, but not of 6. In this reasonable sense of the term divisor in such a connection, the statement quoted becomes intelligible only when modified to read: every integral root of an equation with an integral absolute term is a divisor of that term. But this is not always true. The integral root 6 of $x^2 - \frac{20}{3}x + 4 = 0$ is not a divisor of 4; the root 2 of $x^2 - \frac{1}{2}x - 3 = 0$ is not a divisor of -3. The correct theorem is that next stated.

5. Integral Roots. *For an equation all of whose coefficients are integers, that of the highest power of the variable being unity, any integral root is a divisor of the constant term.*

In certain texts, we find a correct statement of this theorem, but an erroneous proof. When α_1 and p_n are integers and $\alpha_1\alpha_2 \ldots \alpha_n = \pm p_n$, it is falsely concluded that α_1 is a divisor of p_n. But $12 \cdot 3 \cdot \frac{1}{4} = 9$ and 12 is not a divisor of 9. Also the examples at the end of § 4 show the falsity of this argument and, indeed, of any argument not making use of the hypothesis that all of the coefficients are integers.

A correct proof is very easily given. Let d be an integral root of equation (1), in which now $p_1, \ldots, p_n$ are all integers. Then

$$d^n + p_1 d^{n-1} + p_2 d^{n-2} + \cdots + p_{n-1}d + p_n = 0. \tag{5}$$

Since d obviously divides all of the terms preceding the last term, it must divide p_n.

Hence if there be integral roots of an equation of the specified type, they may be found by testing in turn each positive and negative divisor d of the constant term p_n. The most obvious test is to compute (by the abridgment in Ch. I, § 5) the value of $f(d)$ and note whether or not this value is zero. We may shorten the work very much by various methods, and most by a combination of these methods.

Evidently it is unnecessary to test a value of d beyond the limits of the positive and negative roots.

6. Newton's Method for Integral Roots. Consider an equation (1) with integral coefficients. Let d be an integral root. It is a divisor of p_n and we may set

$$p_n = dq_{n-1}.$$

By removing the factor d from each term of (5), we get

$$d^{n-1} + p_1 d^{n-2} + \cdots + p_{n-2}d + p_{n-1} + q_{n-1} = 0.$$

The left member is divisible by d, and hence

$$p_{n-1} + q_{n-1} = dq_{n-2},$$

where q_{n-2} is an integer. Then

$$d^{n-2} + p_1 d^{n-3} + \cdots + p_{n-3}d + p_{n-2} + q_{n-2} = 0,$$

$$p_{n-2} + q_{n-2} = dq_{n-3},$$

where q_{n-3} is an integer, etc. Conversely, if such a relation holds at each step and if, finally, $1 + q_0$ is zero, then d is a root, and the quotient of $f(x)$ by $x - d$ is

$$x^{n-1} - q_1 x^{n-2} - q_2 x^{n-3} - \cdots - q_{n-2}x - q_{n-1}.$$

Indeed, in the product of the latter by $x - d$, the coefficient of x^{n-t} for $t > 0$ is $dq_{t-1} - q_t$ and this equals p_t by our relations.

COROLLARY. If d is an integral root of an equation $f(x) = x^n + \cdots = 0$ with integral coefficients, the quotient of $f(x)$ by $x - d$ is a polynomial with integral coefficients.

This process is a modification of synthetic division (Ch. X, § 4).

EXAMPLE. $f(x) = x^4 - 9x^3 + 24x^2 - 23x + 15 = 0$. Since evidently there is no negative root, and since 10 is an upper limit to the positive roots, we have only to test the divisors 1, 3, 5 of 15. Now $f(1) = 8$. For $d = 3$, the work is as follows:

$$\begin{array}{ccccc} 1 & -9 & 24 & -23 & 15 \\ \underline{-1} & \underline{6} & \underline{-6} & \underline{5} & \\ 0 & -3 & 18 & -18 & \end{array}$$

Here we have divided 15 by 3 and placed the quotient under -23. Adding, we get -18, whose quotient by 3 is added to 24, etc. Since the last sum is zero, 3 is a root. The quotient has as its coefficients the negatives of the numbers in the second line (see the first line below). We test this quotient for the root 5:

$$\begin{array}{cccc} 1 & -6 & 6 & -5 \\ \underline{-1} & \underline{1} & \underline{-1} & \\ 0 & -5 & 5 & \end{array}$$

Hence 5 is a root and the quotient is $x^2 - x + 1$. The latter does not vanish for $x = \pm 1$. Hence 3 and 5 are the only integral roots and each is a simple root. If we had tested a divisor -3 or 15, not a root, a certain quotient would not be integral and the work would be stopped at that point.

7. Another Method. A divisor d is to be rejected if $d - m$ is not a divisor of $f(m)$, where m is any chosen integer.

For, if d is an integral root of $f(x) = 0$,

$$f(x) \equiv (x - d)\, Q(x),$$

where $Q(x)$ is a polynomial with integral coefficients (§ 6, Cor.). Then $f(m) = (m - d)q$, where q is the integer $Q(m)$.

In the example of § 6, $f(1) = 8$ is not divisible by 14, so that 15 is not an integral root.

Consider the new example

$$f(x) \equiv x^3 - 20\,x^2 + 164\,x - 400 = 0.$$

There is no negative root and 20 is an upper limit to the roots. The positive divisors of 400 less than 20 are 1, 2, 4, 16, 5, 8, 10. The last three are excluded since $f(1) = -255$ is not divisible by 4, 7, or 9. Also 16 is excluded since $f(2) = -144$ is not divisible by 14. Incidentally we have excluded the divisors 1 and 2. The remaining divisor 4 is seen to be a root either by Newton's method or by computing $f(4)$.

In case there are numerous divisors within the limits to the roots, it is usually better not to begin by listing all of the divisors to be tested. For, if a divisor is found to be a root, it is preferable to proceed with the quotient, as was done in the Example in § 6.

EXERCISES

Find all the integral roots of

1. $x^3 - 10\,x^2 + 27\,x - 18 = 0.$
2. $x^4 - 2\,x^3 - 21\,x^2 + 22\,x + 40 = 0.$
3. $x^5 + 47\,x^4 + 423\,x^3 + 140\,x^2 + 1213\,x - 420 = 0.$
4. $x^5 - 34\,x^3 + 29\,x^2 + 212\,x - 300 = 0.$

8. Rational Roots. *Any rational root of an equation with integral coefficients, that of the highest power of the variable being unity, is necessarily an integer.*

Let a/b be a root, where a and b are integers with no common divisor greater than unity. Set $x = a/b$ in (1) and multiply the members of the resulting relation by b^{n-1}. We get

$$\frac{a^n}{b} + p_1 a^{n-1} + p_2 a^{n-2} b + \cdots + p_{n-1} a b^{n-2} + p_n b^{n-1} = 0.$$

All of the terms after the first are integers. Hence b divides a^n. Unless $b = \pm 1$, b has a prime factor which divides a^n and hence also a, contrary to hypothesis. Thus $a/b = \pm a$ is an integral root.

The rational roots of any equation with rational coefficients can now be readily found. If l is the least common denominator of the fractional coefficients, we multiply the members of the equation by l and obtain an equation

$$a_0 y^n + a_1 y^{n-1} + \cdots + a_n = 0,$$

where $a_0, \ldots, a_n$ are integers. Multiply the left member by a_0^{n-1} and set $a_0 y = x$. We obtain an equation (1) with integral coefficients, that of x^n being unity. To any rational root y_1 of the equation in y corresponds a rational root $a_0 y_1$ of (1), which must be an integer, in view of the theorem just proved. Hence we need only find all of the integral roots of the new equation (1) and divide them by a_0 to get all of the rational roots y of the original equation.

Frequently it is sufficient (and of course simpler) to set $ky = x$, where k is a suitable integer less than a_0.

EXERCISES

Find all of the rational roots of

1. $y^4 - \frac{40}{3} y^3 + \frac{130}{3} y^2 - 40 y + 9 = 0$.
2. $6 y^3 - 11 y^2 + 6 y - 1 = 0$.
3. $108 y^3 - 270 y^2 - 42 y + 1 = 0$. [Use $k = 6$.]
4. $32 y^3 - 6 y - 1 = 0$. [Use the least k.]

Form the equation whose roots are the products of 6 by the roots of

5. $x^2 - 2 x - \frac{1}{3} = 0$.
6. $x^3 - \frac{1}{2} x^2 - \frac{1}{3} x + \frac{1}{4} = 0$.

CHAPTER VII

Symmetric Functions

1. Σ-polynomials; Elementary Symmetric Functions. A polynomial in the independent variables $x_1, x_2, \ldots, x_n$ is called *symmetric* in them if it is unaltered by the interchange of any two of the variables. For example,

$$x_1^2 + x_2^2 + x_3^2 + 3\,x_1 + 3\,x_2 + 3\,x_3$$

is a symmetric function of x_1, x_2, x_3. The sum of the first three terms is denoted by Σx_1^2 and the sum of the last three by $3\,\Sigma x_1$. In general, if t is a product of powers of $x_1, \ldots, x_n$, whose exponents are integers $\geqq 0$, Σt denotes the sum of this term t and all of the distinct terms obtained from it by permutations of the variables. Since such a Σ-polynomial Σt is unaltered by every permutation of the variables, it is unaltered in particular by the interchange of any two variables and hence is a symmetric function. For example, if there are three variables α, β, γ,

$$\Sigma\alpha^2\beta^2\gamma = \alpha^2\beta^2\gamma + \alpha^2\gamma^2\beta + \beta^2\gamma^2\alpha,$$

$$\Sigma\alpha^2\beta^3\gamma = \alpha^2\beta^3\gamma + \beta^2\alpha^3\gamma + \alpha^2\gamma^3\beta + \gamma^2\alpha^3\beta + \beta^2\gamma^3\alpha + \gamma^2\beta^3\alpha.$$

Just as in the case of the initial example, any symmetric polynomial is evidently a linear combination of Σ-polynomials with constant coefficients.

The Σ-polynomials, of the first degree in each variable,

$$(1)\qquad E_1 = \Sigma x_1, \quad E_2 = \Sigma x_1x_2, \quad E_3 = \Sigma x_1x_2x_3, \quad \ldots, \quad E_n = x_1x_2 \ldots x_{n-1}x_n$$

are called the *elementary symmetric functions* of $x_1, \ldots, x_n$.

Frequently we shall employ the notation $\alpha_1, \ldots, \alpha_n$ for the independent variables. By Ch. VI, § 1, $\alpha_1, \ldots, \alpha_n$ are the roots of an equation of degree n,

$$(2)\qquad f(x) \equiv x^n + p_1x^{n-1} + p_2x^{n-2} \cdots + p_n = 0,$$

in which $-p_1, p_2, -p_3, \ldots, (-1)^np_n$ equal the elementary symmetric functions of the roots. It is customary to make the latter statement also for an equation whose roots are not independent variables.

But in the latter case it is preferable to say that $-p_1, p_2, \ldots$ equal the elementary symmetric functions formed for the roots, thus indicating that we have in mind the values of certain functions of arbitrary variables $x_1, \ldots, x_n$ for $x_1 = \alpha_1, \ldots, x_n = \alpha_n$. It may happen that the resulting polynomials in $\alpha_1, \ldots, \alpha_n$ are not symmetric in $\alpha_1, \ldots, \alpha_n$. For example, if the three roots are α, β, β, we have $-p_1 = \alpha + 2\beta$, $p_2 = 2\alpha\beta + \beta^2$, $-p_3 = \alpha\beta^2$, which are the values of $x_1 + x_2 + x_3$, etc., but are not themselves symmetric in α, β, β, being altered by the interchange of α and β.

However, this point will give no trouble in the exercises below, since the roots are given distinct notations and may, if it is desired, be regarded as independent variables.

2. Products of Σ-polynomials. It is a fundamental theorem that any symmetric polynomial in the roots is expressible rationally and integrally in terms of $p_1, p_2, \ldots, p_n$ and the coefficients of the symmetric polynomial. To prove this, it suffices to show that any Σ-polynomial is expressible rationally and integrally in terms of the elementary symmetric functions. Postponing the general proof, we shall now treat several special cases and assign others as exercises.

EXAMPLE 1. If α, β, γ are the roots of $x^3 + px^2 + qx + r = 0$,

$$p^2 = (\alpha + \beta + \gamma)^2 = \alpha^2 + \beta^2 + \gamma^2 + 2(\alpha\beta + \alpha\gamma + \beta\gamma) = \Sigma\alpha^2 + 2q,$$

$$\Sigma\alpha^2 = p^2 - 2q, \quad -pq = \Sigma\alpha \cdot \Sigma\alpha\beta = \Sigma\alpha^2\beta + 3\alpha\beta\gamma, \quad \Sigma\alpha^2\beta = 3r - pq,$$

$$\Sigma\alpha^2\beta\gamma = pr, \quad \Sigma\alpha^2\beta^2 = (\Sigma\alpha\beta)^2 - 2\alpha\beta\gamma\Sigma\alpha = q^2 - 2pr.$$

The student should carry out in detail the steps here indicated.

EXAMPLE 2. The student should learn how to express a product like $\Sigma\alpha \cdot \Sigma\alpha\beta$ in Ex. 1 as a sum of Σ-functions without writing out their expansions, since the latter method is very laborious in general. To obtain the types of Σ-functions in the product, it suffices to use a single term (called leader) of one factor, say α. Then if we use any term of $\Sigma\alpha\beta$ which contains α, we get a term of $\Sigma\alpha^2\beta$; while if we use any term not containing α (hence $\beta\gamma$ in this example), we get a term $\alpha\beta\gamma$. It remains to find the coefficients of these Σ-functions $\Sigma\alpha^2\beta$ and $\alpha\beta\gamma$. To get $\alpha^2\beta$, we must take the term α of $\Sigma\alpha$ and the term $\alpha\beta$ of $\Sigma\alpha\beta$, so that $\Sigma\alpha^2\beta$ has the coefficient unity. To get $\alpha\beta\gamma$, we may take α or β or γ from $\Sigma\alpha$ and the complementary factor $\beta\gamma$ or $\alpha\gamma$ or $\alpha\beta$, respectively, from $\Sigma\alpha\beta$. Hence

$$\underset{3}{\Sigma\alpha} \cdot \underset{3}{\Sigma\alpha\beta} = \underset{6}{\Sigma\alpha^2\beta} + 3\alpha\beta\gamma.$$

As a check, we have marked under each Σ the number* of its terms. Then the total number of terms is $3 \times 3 = 6 + 3$.

* Found by the theory of combinations in Algebra, and not by writing out in full the Σ-functions.

EXAMPLE 3. To find the product of the Σ-functions

$$\Sigma\alpha\beta, \quad s = \Sigma\alpha^2\beta,$$

of $\alpha, \beta, \gamma, \delta$, we use the leader $\alpha\beta$ of the first. To obtain the four types of Σ-functions in the product, we first use a term of s containing both α and β; second, a term of s containing α^2 but not β; third, a term with α but with neither α^2 nor β; fourth, a term free of α and β. The respective types are those in

$$\underset{6}{\Sigma\alpha\beta} \cdot \underset{12}{\Sigma\alpha^2\beta} = 1\,\underset{12}{\Sigma\alpha^3\beta^2} + 2\,\underset{12}{\Sigma\alpha^3\beta\gamma} + 2\,\underset{12}{\Sigma\alpha^2\beta\gamma^2} + 3\,\underset{4}{\Sigma\alpha\beta\gamma^2\delta}.$$

The coefficient of any Σ-function on the right is obtained by counting the number of ways its leader can be expressed as a product of terms of the Σ-functions on the left.

The coefficient of $\alpha^2\beta\gamma^2$ is 2 since we must take either $\alpha\beta$ or $\beta\gamma$ from $\Sigma\alpha\beta$ (for, we must take α or γ, since s does not have a term with two exponents equal to 2; while if we take $\alpha\gamma$, the complementary factor $\alpha\beta\gamma$ is not in s). To obtain $\alpha\beta\gamma^2\delta$, we must take a term from s with γ^2 and α or β or δ. The first and second coefficients are evidently correct.

EXERCISES

If $\alpha, \beta, \gamma, \delta$ are the roots of $x^4 + px^3 + qx^2 + rx + s = 0$, find

1. $\Sigma\alpha^2\beta^2$. [Square $\Sigma\alpha\beta$.]
2. $\Sigma\alpha^3\beta$. [Use $\Sigma\alpha^2 \cdot \Sigma\alpha\beta$.]
3. $\Sigma\alpha^4$. [Square $\Sigma\alpha^2$.]

If α, β, γ are the roots of $x^3 + px^2 + qx + r = 0$, find the cubic equation with the roots

4. $\alpha^2, \beta^2, \gamma^2$. 5. $\alpha\beta, \alpha\gamma, \beta\gamma$. 6. $\dfrac{2}{\alpha}, \dfrac{2}{\beta}, \dfrac{2}{\gamma}$.

By multiplying Σx_1 by a suitable Σ-function, express in terms of functions (1)

7. Σx_1^2 (if $n > 1$). 8. $\Sigma x_1^2 x_2$ if ($n > 2$). 9. $\Sigma x_1^2 x_2$ (if $n = 2$).
10. Σx_1^3 (if $n > 2$). 11. Σx_1^3 (if $n = 2$). 12. $\Sigma x_1^2 x_2 x_3$.
13. For equation (2) with $n > 4$, show that

$$\Sigma\alpha_1^2\alpha_2\alpha_3\alpha_4 = -p_1p_4 + 5\,p_5, \quad \Sigma\alpha_1^2\alpha_2^2\alpha_3 = 3\,p_1p_4 - p_2p_3 - 5\,p_5.$$

14. For equation (2) with $n > 5$, show that

$$\Sigma\alpha_1^2\alpha_2^2\alpha_3\alpha_4 = p_2p_4 - 4\,p_1p_5 + 9\,p_6, \quad \Sigma\alpha_1^2\alpha_2^2\alpha_3^2 = p_3^2 - 2\,p_2p_4 + 2\,p_1p_5 - 2\,p_6.$$

3. Fundamental Theorem on Symmetric Functions. *Any polynomial symmetric in $x_1, \ldots, x_n$ equals a polynomial in the elementary symmetric functions $E_1, \ldots, E_n$ of the x's.*

The proof, illustrated in Exs. 1 and 2 of § 4, tells us just what elementary symmetric functions should be multiplied together in seeking the expression for a given symmetric polynomial in terms of the E's and hence perfects the tentative method used in the earlier examples.

It suffices to prove the theorem for any homogeneous symmetric polynomial S, *i.e.*, one expressible as a sum of terms

$$h = ax_1^{k_1}x_2^{k_2} \ . \ . \ . \ x_n^{k_n}$$

of constant total degree $k = k_1 + k_2 + \cdots + k_n$ in the x's. Evidently we may assume that no two terms of S have the same set of exponents $k_1, \ . \ . \ . \ , k_n$ (since such terms may be combined into a single one). We shall say that h is *higher* than the term $bx_1^{l_1}x_2^{l_2} \ . \ . \ . \ x_n^{l_n}$ if $k_1 > l_1$, or if $k_1 = l_1$, $k_2 > l_2$, or if $k_1 = l_1$, $k_2 = l_2$, $k_3 > l_3$, . . . , so that the first one of the differences $k_1 - l_1$, $k_2 - l_2$, $k_3 - l_3$, . . . which is not zero is positive.

If the highest term in another symmetric polynomial S' is

$$h' = a'x_1^{k_1'}x_2^{k_2'} \ . \ . \ . \ x_n^{k_n'},$$

and that of S is h, then the highest term in their product SS' is

$$hh' = aa'x_1^{k_1+k_1'} \ . \ . \ . \ x_n^{k_n+k_n'}.$$

Indeed, suppose that SS' has a term, higher than hh',

$$cx_1^{l_1+l_1'} \ . \ . \ . \ x_n^{l_n+l_n'}, \tag{3}$$

which is either a product of terms

$$t = bx_1^{l_1} \ . \ . \ . \ x_n^{l_n}, \quad t' = b'x_1^{l_1'} \ . \ . \ . \ x_n^{l_n'}$$

of S and S' respectively, or is a sum of such products. Since (3) is higher than hh', the first one of the differences

$$l + l_1' - k_1 - k_1', \ . \ . \ . \ , \ l_n + l_n' - k_n - k_n'$$

which is not zero is positive. But, either all of the differences $l_1 - k_1, \ . \ . \ . \ ,$ $l_n - k_n$ are zero or the first one which is not zero is negative, since h is either identical with t or is higher than t. Likewise for the differences $l_1' - k_1', \ . \ . \ . \ , \ l_n' - k_n'$. We therefore have a contradiction.

It follows at once that the highest term in any product of homogeneous symmetric polynomials is the product of their highest terms. Now the highest terms in $E_1, E_2, E_3, \ . \ . \ . \ , E_n$, given by (1), are

$$x_1, \quad x_1x_2, \quad x_1x_2x_3, \ . \ . \ . \ , \quad x_1x_2 \ . \ . \ . \ x_n,$$

respectively. Hence the highest term in $E_1^{a_1}E_2^{a_2} \ . \ . \ . \ E_n^{a_n}$ is

$$x_1^{a_1+a_2+ \ . \ . \ . \ +a_n}x_2^{a_2+ \ . \ . \ . \ +a_n} \ . \ . \ . \ x_n^{a_n}.$$

We next prove that, in the above highest term h of S,

$$k_1 \geqq k_2 \geqq k_3 \ . \ . \ . \geqq k_n.$$

For, if $k_1 < k_2$, the symmetric polynomial S would contain the term

$$ax_1^{k_2}x_2^{k_1}x_3^{k_3} \ . \ . \ . \ x_n^{k_n},$$

which is higher than h. If $k_2 < k_3$, S would contain the term

$$ax_1^{k_1}x_2^{k_3}x_3^{k_2} \ . \ . \ . \ x_n^{k_n},$$

higher than h, etc.

By the above result, the highest term in

$$\sigma = aE_1^{k_1-k_2}E_2^{k_2-k_3} \ . \ . \ . \ E_{n-1}^{k_{n-1}-k_n}E_n^{k_n}$$

is h. Hence $S_1 = S - \sigma$ is a homogeneous symmetric polynomial of the same total degree k as S and having a highest term h_1 not as high as h. As before, we form a product σ_1 of the E's whose highest term is this h_1. Then $S_2 = S_1 - \sigma_1$ is a homogeneous symmetric polynomial of total degree k and with a highest term h_2 not as high as h_1. We must finally reach a difference $S_t - \sigma_t$ which is identically zero. Indeed, there is only a finite number of products of powers of $x_1, \ . \ . \ . \ , x_n$ of total degree k. Among these are the parts $h', h_1', h_2', \ . \ . \ .$ of $h, h_1, h_2, \ . \ . \ .$ with the coefficients suppressed. Since each h_i is not as high as h_{i-1}, the $h', h_1', h_2', \ . \ . \ .$ are all distinct. Hence there is only a finite number of h_i. Since $S_t - \sigma_t \equiv 0$,

$$S = \sigma + S_1 = \sigma + \sigma_1 + S_2 = \cdots = \sigma + \sigma_1 + \sigma_2 + \cdots + \sigma_t.$$

Hence S is a polynomial in $E_1, E_2, \ . \ . \ . \ , E_n$.

4. At each step of the preceding process, we subtracted a product of the E's multiplied by the coefficient of the highest term of the earlier function. It follows that *any symmetric polynomial equals a rational integral function, with integral coefficients, of the elementary symmetric functions and the coefficients of the given polynomial.*

Corollary. *Any symmetric polynomial with integral coefficients can be expressed as a polynomial in the elementary symmetric functions with integral coefficients.*

Instances of this important Corollary are furnished by the results in all of our earlier examples and in those which follow.

EXAMPLE 1. If $S = \Sigma x_1^2x_2^2x_3$ and $n > 4$, we have

$$\sigma = E_2E_3 = S + 3\,\Sigma x_1^2x_2x_3x_4 + 10\,\Sigma x_1x_2x_3x_4x_5,$$
$$S_1 = S - \sigma = -3\,\Sigma x_1^2x_2x_3x_4 - 10\,\Sigma x_1x_2x_3x_4x_5,$$
$$\sigma_1 = -3\,E_1E_4 = -3\,(\Sigma x_1^2x_2x_3x_4 + 5\,\Sigma x_1x_2x_3x_4x_5),$$
$$S_2 = S_1 - \sigma_1 = 5\,\Sigma x_1x_2x_3x_4x_5 = 5\,E_5,$$
$$S = \sigma + S_1 = \sigma + \sigma_1 + S_2 = E_2E_3 - 3\,E_1E_4 + 5\,E_5.$$

EXAMPLE 2. If $S = \Sigma x_1^3x_2x_3$ and $n > 4$,

$$\sigma = E_1^2E_3 = E_1\,(\Sigma x_1^2x_2x_3 + 4\,\Sigma x_1x_2x_3x_4)$$
$$= \Sigma x_1^3x_2x_3 + 2\,\Sigma x_1^2x_2^2x_3 + 3\,\Sigma x_1^2x_2x_3x_4$$
$$+ 4\,(\Sigma x_1^2x_2x_3x_4 + 5\,\Sigma x_1x_2x_3x_4x_5),$$
$$S_1 = S - \sigma = -2\,\Sigma x_1^2x_2^2x_3 - 7\,\Sigma x_1^2x_2x_3x_4 - 20\,\Sigma x_1x_2x_3x_4x_5.$$

Take $\sigma_1 = -2\,E_2E_3$ and proceed as in Ex. 1.

REMARK. The definition of a Σ-polynomial in § 1 may be extended to Σ-functions in general. For instance if there are three variables α, β, γ,

$$\sum\frac{1}{\alpha} = \frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma}, \qquad \sum\frac{\beta}{\alpha} = \frac{\beta}{\alpha} + \frac{\gamma}{\alpha} + \frac{\alpha}{\beta} + \frac{\gamma}{\beta} + \frac{\alpha}{\gamma} + \frac{\beta}{\gamma}.$$

EXERCISES

If $\alpha, \beta, \gamma, \delta$ are the roots of $x^4 + px^3 + qx^2 + rx + s = 0$,

1. $\displaystyle\sum\frac{1}{\alpha} = \frac{-r}{s}.$

2. $\displaystyle\sum\frac{\beta}{\alpha} = \sum\alpha\cdot\sum\frac{1}{\alpha} - 4 = \frac{pr}{s} - 4.$

3. $\displaystyle\sum\frac{1}{\alpha\beta} = \frac{q}{s}.$

4. $\displaystyle\sum\frac{1}{\alpha^2} = \frac{1}{s^2}(r^2 - 2\,qs).$

5. $\displaystyle\sum\frac{\beta\gamma}{\alpha^2} = \sum\alpha\beta\cdot\sum\frac{1}{\alpha^2} - \sum\frac{\beta}{\alpha} = \frac{1}{s^2}(qr^2 - 2\,q^2s - prs + 4\,s^2).$

6. Find the sums in Exs. 1, 3, 4 from the sum, sum of the products two at a time, and sum of the squares of the roots of

$$1 + py + qy^2 + ry^3 + sy^4 = 0,$$

obtained by replacing x by $1/y$ in the former quartic equation.

7. $\displaystyle\sum_{12}\frac{\beta}{\alpha} = \sum_{4}\frac{\beta+\gamma+\delta}{\alpha} = \sum\frac{-p-\alpha}{\alpha} = -4 - p\sum\frac{1}{\alpha}.$

8. $\displaystyle\sum_{6}\frac{\alpha^2+\beta^2}{\alpha\beta} = \sum_{12}\frac{\beta}{\alpha}.$

9. $\displaystyle\sum\frac{\gamma}{\alpha\beta} = \frac{3\,r - pq}{s}.$

10. $\sum\frac{1}{\alpha}\cdot\sum\frac{\beta}{\alpha}=\sum\frac{\beta}{\alpha^2}+3\sum\frac{1}{\alpha}+2\sum\frac{\gamma}{\alpha\beta}$. 11. $\sum\frac{\beta}{\alpha^2}=\frac{1}{s^2}(rs-pr^2+2\,pqs)$.

12. Prove that the degree in any single x of a homogeneous symmetric polynomial S is the total degree of the equal polynomial in the E's. Hints: First show that no term of S has an exponent $> k_1$, so that the degree of S in any single x is k_1. Next, σ is of total degree k_1 in the E's. Set $h_1 = a'x_1^{k_1'} \ldots$. Then σ_1 is of total degree $k_1'(\leqq k_1)$ in the E's and not every exponent in σ_1 equals the corresponding exponent in σ. Thus σ is not cancelled by $\sigma_1, \sigma_2, \ldots$.

13. Given a polynomial in the E's of total degree d, show that the equal function of the x's is of degree $\leqq d$ in any single root.

5. Sums of Like Powers of the Roots. If $\alpha_1, \ldots, \alpha_n$ are the roots of (2), we write $s_1 = \Sigma\alpha_1$, $s_2 = \Sigma\alpha_1^2$, and, in general,

$$s_k = \Sigma\alpha_1^k = \alpha_1^k + \alpha_2^k + \cdots + \alpha_n^k.$$

The factored form of (2) is

$$f(x) \equiv (x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_n). \tag{4}$$

In this identity in x, we may replace x by $x + h$. Thus

$$f(x+h) \equiv (x+h-\alpha_1)(x+h-\alpha_2)\ldots(x+h-\alpha_n).$$

In the expansion of $f(x + h)$ as a polynomial in h, the coefficient of the first power of h is $f'(x)$, by the definition of the first derivative of $f(x)$ in Ch. I, § 4. In the right member, the coefficient of h is

$$(x-\alpha_2)(x-\alpha_3)\ldots(x-\alpha_n)+\cdots+(x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_{n-1}).$$

Here the first product equals $f(x) \div (x - \alpha_1)$, by (4), etc. Hence

$$f'(x) \equiv \frac{f(x)}{x-\alpha_1}+\frac{f(x)}{x-\alpha_2}+\cdots+\frac{f(x)}{x-\alpha_n}. \tag{5}$$

If α is any root of (2), $f(\alpha) = 0$ and

$$\begin{aligned}\frac{f(x)}{x-\alpha}&=\frac{f(x)-f(\alpha)}{x-\alpha}=\frac{x^n-\alpha^n}{x-\alpha}+p_1\frac{x^{n-1}-\alpha^{n-1}}{x-\alpha}+\cdots+p_{n-1}\frac{x-\alpha}{x-\alpha}\\&=x^{n-1}+\alpha x^{n-2}+\alpha^2x^{n-3}+\cdots+p_1(x^{n-2}+\alpha x^{n-3}+\cdots)\\&\quad+p_2(x^{n-3}+\cdots)+\cdots,\end{aligned}$$

$$\begin{aligned}\frac{f(x)}{x-\alpha}&=x^{n-1}+(\alpha+p_1)x^{n-2}+(\alpha^2+p_1\alpha+p_2)x^{n-3}+\cdots\\&\quad+(\alpha^k+p_1\alpha^{k-1}+p_2\alpha^{k-2}+\cdots+p_{k-1}\alpha+p_k)x^{n-k-1}+\cdots.\end{aligned} \tag{6}$$

Taking α to be $\alpha_1, \ldots, \alpha_n$ in turn and adding the results, we have by (5)

$$\begin{aligned}f'(x)&=nx^{n-1}+(s_1+np_1)x^{n-2}+(s_2+p_1s_1+np_2)x^{n-3}\cdots\\&\quad+(s_k+p_1s_{k-1}+p_2s_{k-2}+\cdots+p_{k-1}s_1+np_k)x^{n-k-1}+\cdots.\end{aligned}$$

By Ch. I., § 4,

$$f'(x) = nx^{n-1} + (n-1)p_1x^{n-2} + (n-2)p_2x^{n-3} + \cdots + (n-k)p_kx^{n-k-1} + \cdots.$$

Since the coefficients of like powers of x are equal, we get

$$(7) \qquad s_1 + p_1 = 0, \quad s_2 + p_1s_1 + 2\,p_2 = 0, \ \ldots,$$

$$s_k + p_1s_{k-1} + p_2s_{k-2} + \cdots + p_{k-1}s_1 + kp_k = 0 \quad (k = 1, 2, \ldots, n-1).$$

We may therefore find in turn $s_1, s_2, \ldots, s_{n-1}$:

$$(8) \qquad s_1 = -p_1, \quad s_2 = p_1^2 - 2\,p_2, \quad s_3 = -p_1^3 + 3\,p_1p_2 - 3\,p_3, \ \ldots.$$

To find s_n, replace x in (2) by $\alpha_1, \ldots, \alpha_n$ in turn and add the resulting equations. We get

$$(9) \qquad s_n + p_1s_{n-1} + p_2s_{n-2} + \cdots + p_{n-1}s_1 + np_n = 0.$$

We may combine (7) and (9) into a single formula:

$$(10) \quad s_k + p_1s_{k-1} + p_2s_{k-2} + \cdots + p_{k-1}s_1 + kp_k = 0 \quad (k = 1, 2, \ldots, n).$$

To derive a formula which shall enable us to compute the s_k for $k > n$, we multiply (2) by x^{k-n}, take $x = \alpha_1, \ldots, x = \alpha_n$ in turn, and add the resulting equations. We get

$$(11) \qquad s_k + p_1s_{k-1} + p_2s_{k-2} + \cdots + p_ns_{k-n} = 0 \qquad (k > n).$$

Relations (10) and (11) are called *Newton's formulæ*. They enable us to express any s_k as a polynomial in $p_1, \ldots, p_n$.

EXERCISES

1. For a cubic equation, $s_4 = p_1^4 - 4\,p_1^2p_2 + 4\,p_1p_3 + 2\,p_2^2$.
2. For an equation of degree $n \geqq 4$, $s_4 = p_1^4 - 4\,p_1^2p_2 + 4\,p_1p_3 + 2\,p_2^2 - 4\,p_4$.
3. If we define $p_{n+1}, p_{n+2}, \ldots$ to be zero, relations (10) hold for every k. Hence if $p_1, p_2, \ldots$ are arbitrary numbers unlimited in number, and if $\sigma_1, \sigma_2, \ldots$ are computed by use of

$$\sigma_k + p_1\sigma_{k-1} + \cdots + p_{k-1}\sigma_1 + kp_k \quad (k = 1, 2, \ldots),$$

σ_k becomes s_k when we take $p_{n+1} = 0$, $p_{n+2} = 0, \ldots$. See Exs. 1, 2.
4. For $x^n - 1 = 0$, $s_k = n$ or 0 according as k is divisible or not by n.

6. Σ-functions Expressed in Terms of the Functions s_k. We have

$$s_as_b = \Sigma\alpha_1^a \cdot \Sigma\alpha_1^b = \Sigma\alpha_1^{a+b} + m\Sigma\alpha_1^a\alpha_2^b,$$

$$(12) \qquad \Sigma\alpha_1^a\alpha_2^b = \frac{1}{m}(s_as_b - s_{a+b}),$$

where $m = 1$ if $a \neq b$, $m = 2$ if $a = b$.

Any Σ-function with a term involving just three roots may be denoted by $\Sigma\alpha_1^a\alpha_2^b\alpha_3^c$, $a \geqq b \geqq c > 0$. If $b > c$,

$$s_a\Sigma\alpha_1^b\alpha_2^c = m\Sigma\alpha_1^a\alpha_2^b\alpha_3^c + \Sigma\alpha_1^{a+b}\alpha_2^c + \Sigma\alpha_1^{a+c}\alpha_2^b,$$

for m as above. Since $a + b > c$, $a + c > b$,

$$m\Sigma\alpha_1^a\alpha_2^b\alpha_3^c = s_a(s_bs_c - s_{b+c}) - (s_{a+b}s_c - s_{a+b+c}) - (s_{a+c}s_b - s_{a+b+c}),$$

$$\Sigma\alpha_1^a\alpha_2^b\alpha_3^c = \frac{1}{m}(s_as_bs_c - s_as_{b+c} - s_bs_{a+c} - s_cs_{a+b} + 2\,s_{a+b+c}) \quad (b > c). \tag{13}$$

But if $b = c$, we have

$$s_a\Sigma\alpha_1^b\alpha_2^b = r\,\Sigma\alpha_1^a\alpha_2^b\alpha_3^b + \Sigma\alpha_1^{a+b}\alpha_2^b,$$

where $r = 1$ if $a > b$, $r = 3$ if $a = b$. Hence

$$\Sigma\alpha_1^a\alpha_2^b\alpha_3^b = \tfrac{1}{2}(s_as_b^2 - s_as_{2b} - 2\,s_bs_{a+b} + 2\,s_{a+2b}) \qquad (a > b), \tag{14}$$

$$\Sigma\alpha_1^a\alpha_2^a\alpha_3^a = \tfrac{1}{6}(s_a^3 - 3\,s_as_{2a} + 2\,s_{3a}). \tag{15}$$

The fact that *any Σ-polynomial can be expressed as a polynomial in the functions* s_k is readily proved by induction. We have

$$s_a\Sigma\alpha_1^b\alpha_2^c\ \ldots\ \alpha_r^g = t\Sigma\alpha_1^a\alpha_2^b\alpha_3^c\ \ldots\ \alpha_{r+1}^g + t_1\Sigma\alpha_1^{a+b}\alpha_2^c\ \ldots\ \alpha_r^g + \cdots + t_r\Sigma\alpha_1^b\alpha_2^c\ \ldots\ \alpha_r^{a+g},$$

where t is a positive integer, and $t_1, \ldots, t_r$ are integers $\geqq 0$ (for example, $t_r = 0$ if $g = b$, since the terms which it multiplies are included in the sum multiplied by t_1). Hence if every $\Sigma\alpha_1^{k_1}\ \ldots\ \alpha_r^{k_r}$ is expressible as a polynomial in the functions s_k, the same is true of every $\Sigma\alpha_1^a\alpha_2^b\ \ldots\ \alpha_{r+1}^g$. But the theorem is true for $r = 1$ (by the definition of s_k). Hence it is true by induction for every r.

EXERCISES

1. Take $a = b$ in (13) and then replace c by a. Hence (14) holds also when $a < b$. Derive this result just as we did (14).

2. Express $\Sigma\,\alpha_1^a\alpha_2^b\alpha_3^c\alpha_4^d$ in terms of the s_k, treating all cases. Why are these formulæ unnecessary if the equation is of degree four?

3. For a quartic equation express the functions

$$\Sigma\alpha_1^2\alpha_2^2, \quad \Sigma\alpha_1^3\alpha_2, \quad \Sigma\alpha_1^2\alpha_2\alpha_3, \quad \Sigma\alpha_1^2\alpha_2^2\alpha_3$$

in terms of the s_k and ultimately in terms of the $p_1, \ldots, p_4$.

7. Since any s_k equals a polynomial in $p_1, \ldots, p_n$(§ 5), the theorem of § 6 shows that any Σ-polynomial (and hence any rational integral symmetric function) of the roots of an equation equals a polynomial in

the coefficients $p_1, \ldots, p_n$ of the equation. Since we may form an equation with arbitrarily assigned roots, we have a new proof of the fundamental theorem on symmetric functions (§ 3).

The method of §§ 5, 6 to express a Σ-polynomial in terms of the coefficients is advantageous when a term of Σ involves only a few distinct roots, but with high exponents, while the method of §§ 2, 3 is preferable when a term of Σ involves a large number of roots with low exponents.

8. Waring's Formula * for s_k in Terms of the Coefficients. We shall first derive this formula by a very brief argument employing infinite series in a complex variable, and later give a longer but more elementary proof.

In (2) and (4) replace x by $1/y$ and multiply each by y^n. Thus

$$(16) \qquad 1 + p_1y + p_2y^2 + \cdots + p_ny^n \equiv (1 - \alpha_1y)(1 - \alpha_2y) \ldots (1-\alpha_ny).$$

Take the natural logarithm of each member, noting that the logarithm of a product equals the sum of the logarithms of the factors, and that

$$\log(1 - z) = -z - \tfrac{1}{2}z^2 - \tfrac{1}{3}z^3 - \cdots - \frac{1}{r}z^r - \cdots = -\sum_{r=1}^{\infty} \frac{1}{r}z^r,$$

if the absolute value of z is < 1. Hence

$$-\sum_{r=1}^{\infty}(-1)^r\frac{1}{r}(p_1y + \cdots + p_ny^n)^r = -\sum_{r=1}^{\infty}\frac{1}{r}(\alpha_1{}^r + \cdots + \alpha_n{}^r)y^r$$

$$\equiv -\sum_{k=1}^{\infty}\frac{1}{k}s_ky^k,$$

if y is sufficiently small in absolute value to ensure the convergence of each of the series used. The coefficient of y^k in $(p_1y + \cdots + p_ny^n)^r$ may be found by the multinomial theorem. Hence, after dividing $r = r_1 + \cdots + r_n$ into the multinomial coefficient, we get

$$(17) \qquad s_k = \sum \frac{(-1)^{r_1+\cdots+r_n}k \cdot (r_1 + \cdots + r_n - 1)!}{r_1!\, r_2! \ldots r_n!} p_1{}^{r_1}p_2{}^{r_2} \ldots p_n{}^{r_n},$$

* Edward Waring, *Misc. Analyt.*, 1762; *Meditationes Algebraicæ*, 1770, p. 225, 3d ed., 1782, pp. 1–4. No hint is given as to how Waring found (17); his proof was in effect by mathematical induction, being a verification that $s_k, s_{k-1}, \ldots, s_1$ satisfy Newton's formulæ.

But (17) had been given earlier by Albert Girard, *Invention nouvelle en l'algèbre* Amsterdam, 1629.

where the sum extends over all sets of integers $r_1, \ldots, r_n$, each $\geqq 0$, for which

$$r_1 + 2r_2 + 3r_3 + \cdots + nr_n = k. \tag{18}$$

Here $r!$ denotes $1 \cdot 2 \cdot 3 \ldots r$ if $r \geqq 1$, and unity if $r = 0$.

9. Elementary Proof of Waring's Formula. Divide each member of (16) into the negative of its derivative; we get

$$\frac{-p_1 - 2p_2y - \cdots - np_ny^{n-1}}{1 + p_1y + \cdots + p_ny^n} \equiv \frac{\alpha_1}{1-\alpha_1 y} + \cdots + \frac{\alpha_n}{1-\alpha_n y}. \tag{19}$$

In the identity

$$\frac{1}{1-Q} \equiv 1 + Q + Q^2 + \cdots + Q^{k-1} + \frac{Q^k}{1-Q}, \tag{20}$$

set $Q = \alpha_g y$ and multiply the resulting terms by α_g. Hence the second member of (19) equals

$$s_1 + s_2y + \cdots + s_ky^{k-1} + \frac{y^k\phi(y)}{1 + p_1y + \cdots + p_ny^n}, \tag{21}$$

the polynomial $\phi(y)$ being introduced in bringing the fractional terms

$$\alpha_1^{k+1}/(1-\alpha_1y),$$

etc., to the common denominator (16).

In (20), we now set $Q = -p_1y - \cdots - p_ny^n$. Thus

$$\frac{1}{1+p_1y+\cdots+p_ny^n} \equiv \sum_{r=0}^{k-1}(-1)^r(p_1y + \cdots + p_ny^n)^r + \frac{y^k\psi(y)}{1+p_1y+\cdots},$$

where $\psi(y)$ is a polynomial. Expanding this rth power by the multinomial theorem, we see that the left member of (19) equals

$$d\sum(-1)^{r_1+\cdots+r_n+1}\frac{(r_1+\cdots+r_n)!}{r_1!\cdots r_n!}p_1^{r_1}\cdots p_n^{r_n}y^{r_1+2r_2+\cdots+nr_n} + E$$

$$(d = p_1 + 2p_2y + \cdots),$$

the sum extending over all integral values $\geqq 0$ of $r_1, r_2, \ldots, r_n$ such that $r_1 + \cdots + r_n < k$, while E is a fraction whose denominator is $1 + p_1y + \cdots$ and whose numerator is the product of y^k by a polynomial in y. In the expansion of the part preceding E, the terms with the factor y^k may be combined with E after they are reduced to the same denominator

as E. The resulting expression* is now of the same general form as (21), so that the coefficient of y^{k-1} must equal s_k. This coefficient is the sum of

$$\sum(-1)^{r_1+\cdots+r_n+1}\frac{(r_1+\cdots+r_n)!}{r_1!\ldots r_n!}p_1^{r_1+1}p_2^{r_2}\ldots p_n^{r_n}$$
$$(r_1+2r_2+\cdots+nr_n=k-1),$$
$$2\sum(-1)^{r_1+\cdots+r_n+1}\frac{(r_1+\cdots+r_n)!}{r_1!\ldots r_n!}p_1^{r_1}p_2^{r_2+1}\ldots p_n^{r_n}$$
$$(r_1+2r_2+\cdots+nr_n=k-2),$$
$$3\sum(-1)^{r_1+\cdots+r_n+1}\frac{(r_1+\cdots+r_n)!}{r_1!\ldots r_n!}p_1^{r_1}p_2^{r_2}p_3^{r_3+1}\ldots p_n^{r_n}$$
$$(r_1+2r_2+\cdots+nr_n=k-3),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$

In the first sum employ the summation index r_1+1 instead of r_1; in the second sum, r_2+1 instead of r_2; etc. We get

$$\sum(-1)^{r_1+\cdots+r_n}\frac{(r_1+\cdots+r_n-1)!}{(r_1-1)!\,r_2!\ldots r_n!}p_1^{r_1}\ldots p_n^{r_n},$$
$$2\sum(-1)^{r_1+\cdots+r_n}\frac{(r_1+\cdots+r_n-1)!}{r_1!\,(r_2-1)!\ldots r_n!}p_1^{r_1}\ldots p_n^{r_n},$$
$$3\sum(-1)^{r_1+\cdots+r_n}\frac{(r_1+\cdots+r_n-1)!}{r_1!\,r_2!\,(r_3-1)!\ldots r_n!}p_1^{r_1}\ldots p_n^{r_n},$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$

where now (18) holds for each sum. Adding these sums, we evidently get the second member of (17).

Example 1. Let $n=3$, $k=4$. Then $r_1+2r_2+3r_3=4$ and

$$(r_1, r_2, r_3) = (4, 0, 0),\quad (2, 1, 0),\quad (1, 0, 1),\quad (0, 2, 0),$$
$$s_4 = 4\left(\frac{3!}{4!}p_1^4-\frac{2!}{2!\,1!}p_1^2p_2+\frac{1!}{1!\,1!}p_1p_3+\frac{1!}{2!}p_2^2\right)$$
$$= p_1^4-4p_1^2p_2+4p_1p_3+2p_2^2.$$

* The difference between it and (21) is an expression of the form (21). Suppose therefore that an expression (21) is identically zero. Taking $y=0$, we get $s_1=0$. The quotient by y is identically zero. Then $s_2=0$, etc.

EXAMPLE 2. Let $n = 2$ and write p for $-p_1$, q for p_2, r for r_2. Then $r_1 = k - 2r$. If κ is the largest integer $\leqq k/2$, the sum of the kth powers of the roots of $x^2 - px + q = 0$ is

$$s_k = \sum_{r=0}^{\kappa} \frac{(-1)^r k \cdot (k-r-1)!}{(k-2r)!\,r!} p^{k-2r} q^r$$
$$= p^k - kp^{k-2}q + \frac{k(k-3)}{1 \cdot 2} p^{k-4}q^2 - \frac{k(k-4)(k-5)}{1 \cdot 2 \cdot 3} p^{k-6}q^3 + \cdots.$$

10.† Certain Equations Solvable by Radicals. Regarding p as a variable and q as a constant, denote the polynomial in the preceding Ex. 2 by $F(p)$. *The equation $F(p) = c$, where c is an arbitrary constant, can be solved by radicals.* Indeed, if x is a particular root of $x^2 - px + q = 0$, the second root is q/x, and

$$s_k = x^k + \left(\frac{q}{x}\right)^k.$$

This expression in x is therefore the result of replacing p by $x + q/x$ in $F(p)$, as shown by the quadratic equation. Hence $F(p) = c$ then becomes

$$x^k + \left(\frac{q}{x}\right)^k = c, \quad x^{2k} - cx^k + q^k = 0.$$

Solving this as a quadratic equation for x^k, we get

$$x^k = \frac{c}{2} \pm \sqrt{\frac{c^2}{4} - q^k}.$$

Since the product of these two expressions is q^k, definite values

$$\rho = \sqrt[k]{\frac{c}{2} + \sqrt{\frac{c^2}{4} - q^k}}, \quad \sigma = \sqrt[k]{\frac{c}{2} - \sqrt{\frac{c^2}{4} - q^k}}$$

can be chosen so that $\rho\sigma = q$. Hence if ϵ be a primitive kth root of unity, the $2k$ values of x can be separated into pairs $\rho\epsilon^m$, $\sigma\epsilon^{k-m}$ $(m = 0, 1, \ldots, k-1)$, such that the product of the two in a pair is $\rho\sigma = q$. Now $x + q/x$ is a value of p. Hence *the k roots p of $F(p) = c$ are*

$$\rho\epsilon^m + \sigma\epsilon^{k-m} \qquad (m = 0, 1, \ldots, k-1).$$

Thus $F(p) = c$ can be solved by making the substitution

$$p = x + \frac{q}{x}.$$

For $k = 3$, the equation is $p^3 - 3\,qp = c$ and the present method becomes that in Ch. III for solving a reduced cubic equation.

EXERCISES

1.† Solve DeMoivre's quintic $p^5 - 5\,qp^3 + 5\,q^2p = c$ for p.
2.† Solve $p^4 - 4\,qp^2 + 2\,q^2 = c$ for p by this method.
3.† Write down a solvable equation of degree 7. Solve it.
4.† Solve $y^5 + 10\,y^3 + 20\,y + 31 = 0$.

11. Polynomials Symmetric in all but One of the Roots. *If P is a polynomial in the roots of an equation $f(x) = 0$ of degree n and if P is symmetric in $n - 1$ of the roots, then P equals a polynomial in the remaining root and the coefficients of P and $f(x)$.*

For example, $P = 3\,\alpha_1 + \alpha_2^2 + \alpha_3^2 + \cdots + \alpha_n^2$ is such a polynomial and

$$P = \Sigma\alpha_1^2 + 3\,\alpha_1 - \alpha_1^2 = p_1^2 - 2\,p_2 + 3\,\alpha_1 - \alpha_1^2.$$

If α is the remaining root, P is symmetric in all of the roots of the equation (6) of degree $n - 1$, whose coefficients are polynomials in $\alpha, p_1, \ldots, p_n$. Hence (§ 3) P equals a polynomial in $\alpha, p_1, \ldots, p_n$ and the coefficients of P.

Example 1. If α, β, γ are the roots of $f(x) \equiv x^3 + px^2 + qx + r = 0$, find

$$\sum\frac{\alpha^2 + \beta^2}{\alpha + \beta} = \frac{\alpha^2 + \beta^2}{\alpha + \beta} + \frac{\alpha^2 + \gamma^2}{\alpha + \gamma} + \frac{\beta^2 + \gamma^2}{\beta + \gamma}.$$

Since $\beta^2 + \gamma^2 = p^2 - 2\,q - \alpha^2, \quad \beta + \gamma = -p - \alpha,$

$$\sum\frac{\alpha^2 + \beta^2}{\alpha + \beta} = \sum\frac{p^2 - 2\,q - \alpha^2}{-p - \alpha} = \sum\left(\alpha - p + \frac{2\,q}{\alpha + p}\right) = -p - 3\,p + 2\,q\sum\frac{1}{\alpha + p}.$$

But $\alpha + p, \beta + p, \gamma + p$ are the roots y_1, y_2, y_3 of the cubic equation obtained from $f(x) = 0$ by setting $x + p = y$, *i.e.*, $x = y - p$. The resulting equation is

$$y^3 - 2\,py^2 + (p^2 + q)y + r - pq = 0.$$

Since we desire the sum of the reciprocals of y_1, y_2, y_3, we set $y = 1/z$ and find the sum of the roots z_1, z_2, z_3 of

$$1 - 2\,pz + (p^2 + q)z^2 + (r - pq)z^3 = 0.$$

Hence

$$\sum\frac{1}{\alpha + p} = \sum\frac{1}{y_1} = \sum z_1 = \frac{p^2 + q}{pq - r}, \qquad \sum\frac{\alpha^2 + \beta^2}{\alpha + \beta} = \frac{2\,q^2 - 2\,p^2q + 4\,pr}{pq - r}.$$

EXAMPLE 2. If $x_1, \ldots, x_n$ are the roots of $f(x) = 0$, find

$$\sum \frac{1}{x_1 + c}.$$

First, $x_1 + c = y_1, \ldots, x_n + c = y_n$ are the roots of

$$f(-c+y) = f(-c) + yf'(-c) + y^2(\ ) + \ldots = 0.$$

Next, $1/y_1 = z_1, \ldots, 1/y_n = z_n$ are the roots of

$$z^n f(-c) + z^{n-1}f'(-c) + z^{n-2}(\ ) + \ldots = 0,$$

obtained by setting $y = 1/z$ in the preceding equation. Hence

$$\sum \frac{1}{x_1 + c} = \sum z_1 = \frac{-f'(-c)}{f(-c)}.$$

EXERCISES

[In Exs. 1–14, α, β, γ are the roots of $f(x) = x^3 + px^2 + qx + r = 0$.]

1. Find $\sum \frac{1}{\alpha + p}$ by means of the last formula.

Using $\beta\gamma + \alpha(\beta + \gamma) = q$, find

2. $\sum \frac{\beta\gamma + \alpha^2}{\beta + \gamma}$.   3. $\sum \frac{3\beta\gamma - 2\alpha^2}{\beta + \gamma - \alpha}$.

4. Why would the use of $\beta\gamma = -r/\alpha$ complicate Exs. 2, 3? Verify

$$\beta\gamma = \frac{-r}{\alpha} = \frac{f(\alpha) - r}{\alpha} = \alpha^2 + p\alpha + q.$$

5. Why would you use $\beta\gamma = -r/\alpha$ in finding $\sum \frac{\beta^2 + \gamma^2}{\beta\gamma}$?

6. Show that the last sum equals $\Sigma(\gamma/\beta)$.

7. Find $\sum (\beta + \gamma)^2$.   8. Find $\sum (\alpha + \beta - \gamma)^3$.   9. Find $\sum \left(\frac{\beta - \gamma}{\beta + \gamma}\right)^2$.

10. Find a necessary and sufficient condition on the coefficients that the roots, in some order, shall be in harmonic progression. If $\frac{1}{\alpha} + \frac{1}{\gamma} = \frac{2}{\beta}$, then $\frac{-3r}{q} - \beta = 0$, and conversely. But

$$f\left(\frac{-3r}{q}\right) = \left(\frac{-3r}{q} - \alpha\right)\left(\frac{-3r}{q} - \beta\right)\left(\frac{-3r}{q} - \gamma\right).$$

11. Find the cubic equation with the roots $\beta\gamma - \frac{1}{\alpha}$, $\alpha\gamma - \frac{1}{\beta}$, $\alpha\beta - \frac{1}{\gamma}$. Hint: since these are $(-r-1)/\alpha$, etc., make the substitution $(-r-1)/x = y$.

Find the substitution which replaces the given cubic equation by one with the roots

12. $\alpha\beta + \alpha\gamma, \quad \alpha\beta + \beta\gamma, \quad \alpha\gamma + \beta\gamma.$

13. $\dfrac{2\alpha - 1}{\beta + \gamma - \alpha}$, etc. 14. $\dfrac{\beta\gamma + 3\alpha^2}{\beta + \gamma - 2\alpha}$, etc.

If $\alpha, \beta, \gamma, \delta$ are the roots of $x^4 + px^3 + qx^2 + rx + s = 0$, find

15. $\sum \dfrac{\beta^2 + \gamma^2 + \delta^2}{\beta + \gamma + \delta}$. 16. $\sum \dfrac{\beta\gamma + \beta\delta + \gamma\delta}{\beta + \gamma + \delta - 3}$.

17. If y_1, y_2, y_3 are the roots of $y^3 + py + q = 0$, the equation with the roots $z_1 = (y_2 - y_3)^2, \quad z_2 = (y_1 - y_3)^2, \quad z_3 = (y_1 - y_2)^2$ is

$$z^3 + 6pz^2 + 9p^2z + 4p^3 + 27q^2 = 0.$$

Hints: since $z_1 = \Sigma y_1^2 - 2y_2y_3 - y_1^2 = -2p + 2q/y_1 - y_1^2$, etc., we set $z = -2p + 2q/y - y^2$. By the given equation, $y^2 + p + q/y = 0$. Thus the desired substitution is $z = -p + 3q/y, \quad y = 3q/(z + p)$.

18. Hence find the discriminant of the reduced cubic equation.

12.† Cauchy's Method for Symmetric Functions. If $x_1, \ldots, x_n$ are the roots of (2), any polynomial P in $x_1, \ldots, x_n$ can be expressed as a polynomial in $x_2, \ldots, x_n, p_1, \ldots, p_n$, in every term of which the exponent of x_2 is less than 2, the exponent of x_3 less than 3, . . . , the exponent of x_n less than n. To this end, we first eliminate x_1 by using $\Sigma x_1 = -p_1$. Then we eliminate $x_2^k (k \geqq 2)$ by using the quadratic equation satisfied by x_2 and having as coefficients polynomials in $x_3, \ldots, x_n$. This quadratic may be obtained by dividing $f(x)$ by $(x - x_3) \ldots (x - x_n)$, or by noting that

$$x_1 + x_2 = -p_1 - x_3 - \cdots - x_n,$$

$$x_1x_2 = p_2 - (x_1 + x_2)(x_3 + \cdots + x_n) - x_3x_4 - \cdots - x_{n-1}x_n.$$

Next, we eliminate $x_3^k (k \geqq 3)$ by using the cubic equation obtained by dividing $f(x)$ by $(x - x_4) \ldots (x - x_n)$. Finally, we eliminate $x_n^k (k \geqq n)$ by using $f(x_n) = 0$.

EXAMPLE. To compute by this method the discriminant

$$\Delta = (x_1 - x_2)^2(x_1 - x_3)^2(x_2 - x_3)^2$$

of $f(x) \equiv x^3 + px + q = 0$, we note that x_1 and x_2 are the roots of

$$\frac{f(x)}{x - x_3} = Q(x) \equiv x^2 + xx_3 + x_3^2 + p = 0.$$

Since $\Sigma x_1 = 0$,

$$(x_1 - x_2)^2 = (-2x_2 - x_3)^2 = 4Q(x_2) - 3x_3^2 - 4p = -3x_3^2 - 4p,$$

$$(x_3 - x_1)(x_3 - x_2) = Q(x_3) = 3x_3^2 + p,$$

$$\Delta = (-3x_3^2 - 4p)(3x_3^2 + p)^2 = -27(x_3^3 + px_3)^2 - 4p^3,$$

$$\Delta = -27q^2 - 4p^3.$$

We can now easily prove the fundamental theorem of § 3: if P is symmetric in $x_1, \ldots, x_n$, it equals a polynomial in $p_1, \ldots, p_n$. For, $P = A + Bx_2$, where neither A nor B involves x_1 or x_2. Since P is unaltered when x_1 and x_2 are interchanged,

$$A + Bx_2 = A + Bx_1.$$

If $x_1 \neq x_2$, then $B = 0$; and, by continuity, $B = 0$ even when $x_1 = x_2$. Hence

$$P = C + Dx_3 + Ex_3^2,$$

where C, D, E do not involve x_1, x_2 or x_3. Since P is unaltered when x_3 and x_1 are interchanged, or when x_3 and x_2 are interchanged, the equation

$$0 = C - P + Dy + Ey^2$$

has the three roots x_1, x_2, x_3. Hence if x_1, x_2, x_3 are distinct, $D = E = 0$, $P = C$, and by continuity these relations hold also if two or all three of these x's are equal, so that P is free of x_1, x_2, x_3. Similarly, P can be reduced to a form which is free of each x_i.

13.† Tschirnhausen Transformation. We can eliminate x between (2) and

$$X = u_0 + u_1 x + u_2 x^2 + \cdots + u_{n-1} x^{n-1} \tag{22}$$

and obtain an equation in X of degree n. First, from the expressions for $X^2, X^3, \ldots$, we eliminate $x^n, x^{n+1}, \ldots$ by use of (2) and get

$$\begin{cases} X^2 = u_{20} + u_{21}x + u_{22}x^2 + \cdots + u_{2\,n-1}x^{n-1}, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ X^n = u_{n\,0} + u_{n\,1}x + u_{n\,2}x^2 + \cdots + u_{n\,n-1}x^{n-1}, \end{cases} \tag{23}$$

where the u_{ij} are polynomials in $u_0, \ldots, u_{n-1}$ and the coefficients of (2).

In any one of these equations (23) we set $x = x_1, \ldots, x = x_n$ in turn and add the resulting relations. If $X_1, \ldots, X_n$ are the values of X for $x = x_1, \ldots, x = x_n$, set

$$s_k = \Sigma x_1^k, \quad S_k = \Sigma X_1^k.$$

Then

$$\begin{cases} S_1 = nu_0 + u_1 s_1 + u_2 s_2 + \cdots + u_{n-1} s_{n-1}, \\ S_2 = nu_{20} + u_{21} s_1 + u_{22} s_2 + \cdots + u_{2\,n-1} s_{n-1}, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ S_n = nu_{n\,0} + u_{n\,1} s_1 + u_{n\,2} s_2 + \cdots + u_{n\,n-1} s_{n-1}. \end{cases} \tag{24}$$

Since the elementary symmetric functions of $X_1, \ldots, X_n$ are expressible in terms of $S_1, S_2, \ldots, S_n$ (§ 6), we can find the coefficients of the equation having the roots $X_1, \ldots, X_n$:

$$X^n + P_1X^{n-1} + P_2X^{n-2} + \cdots + P_n = 0. \tag{25}$$

Another method of forming this equation is given in Ch. XII, § 9.

If we seek values of $u_0, u_1, \ldots, u_{n-1}$, such that $P_1, P_2, \ldots, P_k$ shall all vanish and therefore $S_1 = S_2 = \cdots = S_k = 0$, by Newton's identities (7), we have only to satisfy a system of k equations [see (24)] homogeneous in $u_0, \ldots, u_{n-1}$ and of degrees $1, 2, \ldots, k$, respectively. In particular, $S_1 = 0$ enables us to express u_0 in terms of $u_1, \ldots,$ so that

$$X = u_1\left(x - \frac{s_1}{n}\right) + u_2\left(x^2 - \frac{s_2}{n}\right) + \cdots + u_{n-1}\left(x^{n-1} - \frac{s_{n-1}}{n}\right). \tag{26}$$

EXAMPLE. For $n = 3$, $X = u_1(x - \frac{1}{3}s_1) + u_2(x^2 - \frac{1}{3}s_2)$,

$$S_2 = \Sigma X_1^2 = u_1^2(s_2 - \tfrac{1}{3}s_1^2) + 2\,u_1u_2(s_3 - \tfrac{1}{3}s_1s_2) + u_2^2(s_4 - \tfrac{1}{3}s_2^2).$$

Thus $S_2 = 0$ gives

$$(3\,s_2 - s_1^2)\,u_1 = (s_1s_2 - 3\,s_3 + \sqrt{-3\,\Delta})\,u_2,$$

$$\Delta \equiv s_0s_2s_4 + 2\,s_1s_2s_3 - s_0s_3^2 - s_1^2s_4 - s_2^3.$$

Hence the cubic equation is reduced to $X^3 + P_3 = 0$ by the substitution

$$X = (s_1s_2 - 3\,s_3 + \sqrt{-3\,\Delta})(3\,x - s_1) + (3\,s_2 - s_1^2)(3\,x^2 - s_2).$$

By Ex. 6, p. 158, Δ is the discriminant of the cubic equation.

EXERCISES

1.† For $n = 4$, take $u_3 = 0$ in (26) and find the cubic equation for u_1/u_2 which results from $P_3 = 0$ (*i.e.*, $S_3 = 0$, since $S_1 = 0$). The new quartic equation $X^4 + P_2X^2 + P_4 = 0$ may be solved in terms of square roots.

2.† For $n = 5$, the condition for $S_2 = 0$ is that a certain quadratic form q in $u_1, \ldots, u_4$ shall vanish. Now q can be expressed as a sum of the squares of four linear functions L_j of $u_1, \ldots, u_4$. Taking $L_1 = iL_2$, $L_3 = iL_4$, where $i^2 = -1$, we have $S_2 = 0$. By means of the resulting two linear relations between $u_1, \ldots, u_4$, we may express S_3 as a cubic function of u_1, u_2, for example. We must therefore solve a cubic equation in u_1/u_2 to find the u's making also $S_3 = 0$. The new quintic equation is $X^5 + P_4X + P_5 = 0$. If $P_4 \neq 0$, set $X = y\sqrt[4]{P_4}$. Then $y^5 + y + c = 0$. (Bring, 1786; Jerrard, 1834.)

CHAPTER VIII

Reciprocal Equations. Construction of Regular Polygons. Trisection of an Angle

1. For certain types of equations, such as reciprocal and binomial equations, there exist simple relations between the roots, and these relations materially simplify the discussion of the equations.

An equation is called a *reciprocal equation* if the reciprocal of each root is also a root. Apart from possible roots 1 and -1, each of which is its own reciprocal, the roots are in pairs reciprocals of each other.

For example, the equation

$$f(x) = (x-1)(x^2 - \tfrac{5}{2}x + 1) = 0$$

is a reciprocal equation having the roots 1, 2, $\frac{1}{2}$. If we replace x by $1/x$ and multiply the resulting function by x^3, we get $-f(x)$. Here (1) holds for $n = 3$ and for the minus sign.

In general, if

$$f(x) \equiv x^n + \cdots + c = 0$$

is a reciprocal equation, no root is zero, so that $c \neq 0$. If r is any root of $f(x) = 0$, $1/r$ is a root of $f(1/x) = 0$, and hence of

$$x^n f\left(\frac{1}{x}\right) \equiv 1 + \cdots + cx^n = 0.$$

Since the former is a reciprocal equation, it has the root $1/r$. Hence any root of the former equation is a root of the new equation. Thus, by (1) and (2) of Ch. VI, the left member of the latter is the product of $f(x)$ by c. Then, by the constant terms, $1 = c^2$. Hence $c = \pm 1$ and

$$x^n f\left(\frac{1}{x}\right) \equiv \pm f(x). \tag{1}$$

Thus if $p_i x^{n-i}$ is a term of $f(x)$, also $\pm p_i x^i$ is a term. Hence

$$f(x) \equiv x^n \pm 1 + p_1(x^{n-1} \pm x) + p_2(x^{n-2} \pm x^2) + \cdots . \tag{2}$$

If n is *odd*, $n = 2t + 1$, the final term is

$$p_t(x^{t+1} \pm x^t),$$

and $x \pm 1$ is a factor of $f(x)$. In view of (1), the quotient

$$Q(x) \equiv \frac{f(x)}{x \pm 1}$$

has the property that

$$x^{n-1}Q\left(\frac{1}{x}\right) \equiv Q(x).$$

Hence $Q(x) = 0$ is a reciprocal equation of the type

(3) $$x^{2t} + 1 + c_1(x^{2t-1} + x) + c_2(x^{2t-2} + x^2) + \cdots + c_t x^t = 0.$$

Indeed, the highest power x^{2t} of x has the coefficient unity and the constant term is unity, so that it is of the form (2) with the upper signs.

If n is *even*, $n = 2t$, and if the upper sign holds in (1), we have just seen that (2) is of the form (3). Next, let the lower sign hold in (1). Then $p_t = 0$, since a term $p_t x^t$ would imply a term $-p_t x^t$. The final term in (2) is therefore

$$p_{t-1}(x^{t+1} - x^{t-1}).$$

Hence $f(x)$ has the factor $x^2 - 1$. As before, the quotient is of the form (3).

In each case we have been led to a reciprocal equation of type (3). *The solution of the latter may be reduced to the solution of an equation of degree t and certain quadratic equations.* To prove this, divide the terms of (3) by x^t. Then

(4) $$\left(x^t + \frac{1}{x^t}\right) + c_1\left(x^{t-1} + \frac{1}{x^{t-1}}\right) + c_2\left(x^{t-2} + \frac{1}{x^{t-2}}\right) + \cdots + c_{t-1}\left(x + \frac{1}{x}\right) + c_t = 0.$$

To reduce this to an equation of degree t, we set

(5) $$x + \frac{1}{x} = z.$$

Then

$$x^2 + \frac{1}{x^2} = z^2 - 2, \quad x^3 + \frac{1}{x^3} = z^3 - 3z, \ \ldots,$$

while the general binomial in (4) can be computed from

(6) $$x^k + \frac{1}{x^k} = z\left(x^{k-1} + \frac{1}{x^{k-1}}\right) - \left(x^{k-2} + \frac{1}{x^{k-2}}\right).$$

For example,

$$x^4 + \frac{1}{x^4} = z(z^3 - 3z) - (z^2 - 2) = z^4 - 4z^2 + 2.$$

However, we can obtain an explicit expression for $x^k + 1/x^k$ by noting that it is the sum of the kth powers of the roots x, $1/x$ of

$$y^2 - \left(x + \frac{1}{x}\right)y + x \cdot \frac{1}{x} \equiv y^2 - zy + 1 = 0.$$

The sum of the kth powers of the roots of $y^2 - py + q = 0$ was found in Ex. 2, p. 75. Taking $p = z$, $q = 1$, we have

$$(7)\quad x^k + \frac{1}{x^k} = z^k - kz^{k-2} + \frac{k(k-3)}{1 \cdot 2} z^{k-4} - \frac{k(k-4)(k-5)}{1 \cdot 2 \cdot 3} z^{k-6} + \cdots$$
$$+ (-1)^r \frac{k(k-r-1)(k-r-2) \ldots (k-2r+1)}{1 \cdot 2 \cdot 3 \ldots r} z^{k-2r} + \cdots.$$

Hence (4) becomes an equation of degree t in z. From each root z we obtain two roots x of (3), which are reciprocals of each other, by solving the quadratic equation $x^2 - zx + 1 = 0$, equivalent to (5).

EXAMPLE. Solve $x^5 - 5x^4 + 9x^3 - 9x^2 + 5x - 1 = 0$. Dividing by $x - 1$, we get $x^4 - 4x^3 + 5x^2 - 4x + 1 = 0$. Thus

$$x^2 + \frac{1}{x^2} - 4\left(x + \frac{1}{x}\right) + 5 = 0, \quad z^2 - 4z + 3 = 0, \quad z = 1 \text{ or } 3.$$

For $z = 1$, $x^2 - x + 1 = 0$, $x = \frac{1}{2}\left(1 \pm \sqrt{-3}\right)$. For $z = 3$, $x^2 - 3x + 1 = 0$, $x = \frac{1}{2}\left(3 \pm \sqrt{5}\right)$. These with $x = 1$ give the five roots.

EXERCISES

Solve by radicals the reciprocal equations

1. $x^5 - 7x^4 + x^3 - x^2 + 7x - 1 = 0$.
2. $x^5 = 1$.
3. $x^6 = 1$.
4. $x^5 + 1 = 0$.
5. Find the z-cubic for $x^7 = 1$.
6. Find the z-quintic for $x^{11} = 1$.
7. The z-quartic for $x^9 = 1$ is $z^4 + z^3 - 3z^2 - 2z + 1 = 0$. It has the root -1 since the z-equation for $x^3 = 1$ is $z + 1 = 0$. Verify that, on removing the factor $z + 1$ from the quartic, we get the z-cubic $z^3 - 3z + 1 = 0$ for $(x^9 - 1)/(x^3 - 1) = 0$.
8. What are the trigonometric representations of the roots of the z-equations in Exs. 5 and 6? Hint: if $x = \cos\theta + i \sin\theta$, $1/x = \cos\theta - i\sin\theta$.

2. Binomial Reciprocal Equations. A reciprocal equation with only two terms is of the form $x^n \pm 1 = 0$. Its roots were expressed in terms of trigonometric functions in Ch. II. But now we wish to use only algebraic methods.* We might proceed as in § 1, first ** removing the factor $x \pm 1$ (if n is odd) or $x^2 - 1$ (if n is even and the lower sign holds), and then applying substitution (5) to obtain the z-equation. Except for special values of n (as those in Exs. 2–6, § 1), there is a more effective method, leading to auxiliary equations of lower degree than the z-equation. For instance, it will be shown that $x^{17} - 1 = 0$ can be solved in terms of square roots; it is only a waste of effort to form the z-equation of degree 8.

3. The new method will first be illustrated for $x^7 - 1 = 0$ since it then differs only in form from the earlier method of treating reciprocal equations. Removing the factor $x - 1$, we have

$$x^6 + x^5 + x^4 + x^3 + x^2 + x + 1 = 0. \tag{8}$$

If r is a particular root of (8), its six roots are (Ch. II, § 13),

$$r,\ r^2,\ r^3,\ r^4,\ r^5,\ r^6. \tag{9}$$

By the substitution (5), we obtain the cubic equation

$$z^3 + z^2 - 2z - 1 = 0, \tag{10}$$

whose roots are therefore

$$z_1 = r + \frac{1}{r} = r + r^6, \quad z_2 = r^2 + \frac{1}{r^2} = r^2 + r^5, \quad z_3 = r^3 + \frac{1}{r^3} = r^3 + r^4. \tag{11}$$

The new method consists in starting with these sums of pairs of the six roots and forming the cubic equation having these sums as its roots. Since r is a root of (8),

$$\Sigma z_1 = r + r^2 + \cdots + r^6 = -1, \quad \Sigma z_1 z_2 = 2\,(r + \cdots + r^6) = -2,$$
$$z_1 z_2 z_3 = 2 + r + \cdots + r^6 = 1.$$

Hence z_1, z_2, z_3 are the roots of (10). If a root z_1 be found, we can obtain r from the quadratic equation $r^2 - z_1 r + 1 = 0$.

* It is an important fact, not proved or used here, that $x^n \pm 1 = 0$ is solvable by radicals, namely, by a finite number of applications of the operation *extraction of a single root of a known number*. *Cf.* Dickson, *Introduction to the Theory of Algebraic Equations*, John Wiley & Sons, pp. 77, 78. Note that it suffices to treat the case n prime, since $x^{pq} = A$ is equivalent to the chain of equations $y^q = A$, $x^p = y$.

** If $n = pq$, we may remove the factors $x^p \pm 1$ if p is odd. See Ex. 7, §1.

We can, however, find r by solving first a quadratic equation and afterwards a cubic equation. To this end, set

$$(12) \qquad y_1 = r + r^2 + r^4, \qquad y_2 = r^3 + r^6 + r^5.$$

Then

$$y_1 + y_2 = -1, \qquad y_1 y_2 = 3 + r + \cdots + r^6 = 2,$$

so that y_1 and y_2 are the roots of

$$y^2 + y + 2 = 0.$$

Then r, r^2, r^4 are seen to be the roots of

$$\rho^3 - y_1\rho^2 + y_2\rho - 1 = 0.$$

4.† The Periods. We now explain the principle discovered by Gauss by which we select the pairs from (9) to form the *periods* z_1, z_2, z_3 in (11), and the triples to form the periods y_1, y_2 in (12). To this end we seek an integer g such that the six roots (9) can be arranged in the order

$$(13) \qquad r,\ r^g,\ r^{g^2},\ r^{g^3},\ r^{g^4},\ r^{g^5},$$

each term being the gth power of its predecessor. The choice $g = 2$ is not permissible, since the fourth term would then be $r^8 = r$. But we may take $g = 3$, and the desired order is

$$(14) \qquad r,\ r^3,\ r^2,\ r^6,\ r^4,\ r^5,$$

each term being the cube of its predecessor. To form the two periods y_1 and y_2, each of three terms, we take alternate terms of (14). To form the three periods z_1, z_2, z_3, each of two terms, we take any one of the first three terms (as r^3) and the third term after it (then r^4).

5.† Solution of $x^{17} = 1$ by Square Roots. Let r be a root $\neq 1$. Then

$$\frac{r^{17} - 1}{r - 1} = r^{16} + r^{15} + \cdots + r + 1 = 0.$$

As in § 4, we may take $g = 3$ and arrange the roots, $r, \ldots, r^{16}$ so that each is the cube of its predecessor:

$$r,\ r^3,\ r^9,\ r^{10},\ r^{13},\ r^5,\ r^{15},\ r^{11},\ r^{16},\ r^{14},\ r^8,\ r^7,\ r^4,\ r^{12},\ r^2,\ r^6.$$

Taking alternate terms, we form the 2 periods each of 8 terms:

$$y_1 = r + r^9 + r^{13} + r^{15} + r^{16} + r^8 + r^4 + r^2,$$

$$y_2 = r^3 + r^{10} + r^5 + r^{11} + r^{14} + r^7 + r^{12} + r^6.$$

Hence $y_1 + y_2 = -1$. We find that $y_1 y_2 = 4(r + \cdots + r^{16}) = -4$. Thus

$$(15) \qquad y_1,\ y_2 \quad \text{satisfy} \quad y^2 + y - 4 = 0.$$

Taking alternate terms in y_1, we form the two periods

$$z_1 = r + r^{13} + r^{16} + r^4, \quad z_2 = r^9 + r^{15} + r^8 + r^2.$$

Taking alternate terms in y_2, we form the two periods

$$w_1 = r^3 + r^5 + r^{14} + r^{12}, \quad w_2 = r^{10} + r^{11} + r^7 + r^6.$$

Thus $z_1 + z_2 = y_1$, $w_1 + w_2 = y_2$. We find that $z_1 z_2 = w_1 w_2 = -1$. Hence

$$(16) \qquad z_1,\ z_2 \quad \text{satisfy} \quad z^2 - y_1 z - 1 = 0,$$

$$(17) \qquad w_1,\ w_2 \quad \text{satisfy} \quad w^2 - y_2 w - 1 = 0.$$

Taking alternate terms in z_1, we have the periods

$$v_1 = r + r^{16}, \quad v_2 = r^{13} + r^4.$$

Now, $v_1 + v_2 = z_1$, $v_1 v_2 = w_1$. Hence

$$(18) \qquad v_1,\ v_2 \quad \text{satisfy} \quad v^2 - z_1 v + w_1 = 0,$$

$$(19) \qquad r,\ r^{16} \quad \text{satisfy} \quad \rho^2 - v_1 \rho + 1 = 0.$$

Hence we can find r by solving a series of quadratic equations. Which of the sixteen values of r we shall thus obtain depends upon which root of (15) is called y_1 and which y_2, and similarly in (16)–(19). We shall now show what choice is to be made in each such case in order that we shall finally get the value of the particular root

$$r = \cos\frac{2\pi}{17} + i \sin\frac{2\pi}{17}.$$

Then

$$\frac{1}{r} = \cos\frac{2\pi}{17} - i\sin\frac{2\pi}{17}, \quad v_1 = r + \frac{1}{r} = 2\cos\frac{2\pi}{17},$$

$$r^4 = \cos\frac{8\pi}{17} + i\sin\frac{8\pi}{17}, \quad v_2 = r^4 + \frac{1}{r^4} = 2\cos\frac{8\pi}{17}.$$

Hence $v_1 > v_2 > 0$, and therefore $z_1 > 0$. Similarly,

$$w_1 = r^3 + \frac{1}{r^3} + r^5 + \frac{1}{r^5} = 2\cos\frac{6\pi}{17} + 2\cos\frac{10\pi}{17} = 2\cos\frac{6\pi}{17} - 2\cos\frac{7\pi}{17} > 0,$$

$$y_2 = 2\cos\frac{6\pi}{17} + 2\cos\frac{10\pi}{17} + 2\cos\frac{12\pi}{17} + 2\cos\frac{14\pi}{17} < 0,$$

since only the first cosine in y_2 is positive and it is numerically less than the third. But $y_1 y_2 = -4$. Hence $y_1 > 0$. Thus (15)–(17) give

$$y_1 = \tfrac{1}{2}(\sqrt{17} - 1), \qquad y_2 = \tfrac{1}{2}(-\sqrt{17} - 1),$$

$$z_1 = \tfrac{1}{2}y_1 + \sqrt{1 + \tfrac{1}{4}y_1^2}, \qquad w_1 = \tfrac{1}{2}y_2 + \sqrt{1 + \tfrac{1}{4}y_2^2}.$$

We now have the coefficients of (18) and know that $v_1 > v_2 > 0$. These results are sufficient for the next problem. Of course, we could go on and obtain the explicit expression for v_1 and that for r in terms of square roots.

6.† Construction of a Regular Polygon of 17 Sides. In a circle of radius unity, construct two perpendicular diameters AB, CD, and draw

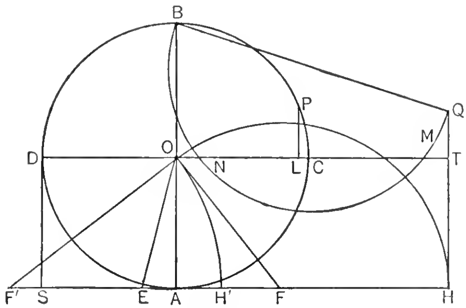


Fig. 20

tangents at A, D, which intersect at S (Fig. 20). Find the point E in AS for which $AE = \frac{1}{4}AS$, by means of two bisections. Then

$$AE = \tfrac{1}{4}, \quad OE = \tfrac{1}{4}\sqrt{17}.$$

Let the circle with center E and radius OE cut AS at F and F'. Then

$$AF = EF - EA = OE - \tfrac{1}{4} = \tfrac{1}{2}y_1,$$
$$AF' = EF' + EA = OE + \tfrac{1}{4} = -\tfrac{1}{2}y_2,$$
$$OF = \sqrt{OA^2 + AF^2} = \sqrt{1 + \tfrac{1}{4}y_1^2}, \quad OF' = \sqrt{1 + \tfrac{1}{4}y_2^2}.$$

Let the circle with center F and radius FO cut AS at H, outside of $F'F$; that with center F' and radius $F'O$ cut AS at H' between F' and F. Then

$$AH = AF + FH = AF + OF = \tfrac{1}{2}y_1 + \sqrt{1 + \tfrac{1}{4}y_1^2} = z_1,$$
$$AH' = F'H' - F'A = OF' - AF' = w_1.$$

It remains to construct the roots of equation (18). This will be done as in Ch. I, § 16. Draw HTQ parallel to AO and intersecting OC produced at T. Make $TQ = AH'$. Draw a circle having as diameter the line BQ joining $B = (0, 1)$ with $Q = (z_1, w_1)$. The abscissas ON and OM of the intersections of this circle with the x-axis OT are the roots of (18). Hence the larger root v_1 is $OM = 2 \cos 2\pi/17$.

Let the perpendicular bisector LP of OM cut the initial circle of unit radius at P. Then

$$\cos LOP = OL = \cos \frac{2\pi}{17}, \quad LOP = \frac{2\pi}{17}.$$

Hence the chord CP is a side of the inscribed regular polygon of 17 sides, constructed with ruler and compasses.

EXERCISES

1. For $n = 5$, $g = 2$, the periods are $r + r^4$, $r^2 + r^3$. Show that they are the roots of the z-quadratic obtained in Ex. 2, p. 83.

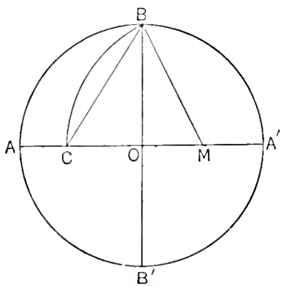


Fig. 21

2.† For $n = 13$, find the least g, form the three periods each of four terms, and find the cubic having them as roots.

3. For $n = 5$, Ex. 1 gives $r + r^4 = 2 \cos 2\pi/5 = \frac{1}{2}(\sqrt{5} - 1)$. In a circle of radius unity and center O draw two perpendicular diameters AOA', BOB'. With the middle point M of OA' as center and radius MB draw a circle cutting OA at C (Fig. 21). Show that OC and BC are the sides s_{10} and s_5 of the inscribed regular decagon and pentagon respectively. Hints:

$$MB = \tfrac{1}{2}\sqrt{5}, \quad OC = \tfrac{1}{2}(\sqrt{5} - 1), \quad BC = \sqrt{1 + \overline{OC}^2} = \tfrac{1}{2}\sqrt{10 - 2\sqrt{5}},$$

$$s_{10} = 2 \sin 18° = 2 \cos \frac{2\pi}{5} = OC,$$

$$s_5^2 = (2 \sin 36°)^2 = 2\left(1 - \cos \frac{2\pi}{5}\right) = \tfrac{1}{4}(10 - 2\sqrt{5}), \quad s_5 = BC.$$

7.† Regular Polygon of n Sides. If n be a prime such that $n - 1$ is a power 2^h of 2 (as is the case when $n = 3, 5, 17$), the $n - 1$ imaginary nth roots of unity can be separated into 2 sets each of 2^{h-1} roots, each of these sets subdivided into 2 sets each of 2^{h-2} roots, etc., until we reach the

sets r, $1/r$ and r^2, $1/r^2$, etc., and in fact * in such a manner that we have a series of quadratic equations, the coefficients of any one of which depend only upon the roots of quadratic equations preceding it in the series. Note that this was the case for $n = 17$ (§ 5) and for $n = 5$. It is in this manner that it can be proved that the roots of $x^n = 1$ can be found in terms of square roots, so that a regular polygon of n sides can be inscribed by ruler and compasses, provided n be a prime of the form $2^h + 1$.

If n be a product of distinct primes of this form, or 2^k times such a product (for example, $n = 15$, 30 or 6), or if $n = 2^m$ $(m > 1)$, it follows readily that we can inscribe by ruler and compasses a regular polygon of n sides. But this is impossible for other values of n. This impossibility will be proved for $n = 7$ and $n = 9$, the method of proof being applicable to the general case.

8. Regular Polygons of 7 and 9 Sides; Trisection of an Angle. For brevity we shall occasionally use the term "construct" for "construct by ruler and compasses." If it were possible to construct a regular polygon of 7 sides and hence angle $2\pi/7$, we could construct a line of length $2\cos 2\pi/7$, the base of a right-angled triangle whose hypotenuse is of length 2 and one of whose acute angles is $2\pi/7$. Set

$$r = \cos\frac{2\pi}{7} + i\sin\frac{2\pi}{7}.$$

Then

$$\frac{1}{r} = \cos\frac{2\pi}{7} - i\sin\frac{2\pi}{7}, \quad r + \frac{1}{r} = 2\cos\frac{2\pi}{7}.$$

Hence $2\cos 2\pi/7$ is a root of the cubic equation (10). This equation has no rational root. For, if it had a rational root, it would have (Ch. VI, § 8, § 5) an integral root which is a divisor of the constant term -1, whereas neither $+1$ nor -1 is a root. Hence we shall know that it is impossible to construct a regular polygon of 7 sides by ruler and compasses as soon as we have proved (§ 10) the next theorem.

* See the author's article "Constructions with ruler and compasses; regular polygons," in *Monographs on Topics of Modern Mathematics*, edited by J. W. A. Young, Longmans, Green and Co., New York, 1911, p. 374. In addition to the references there given (p. 386), mention should be made of the book by Klein, *Elementarmathematik vom Höheren Standpunkte aus*, Leipzig, 1908, vol. 1, p. 125; and ed. 2, 1911.

Theorem. *It is not possible to construct by ruler and compasses a line whose length is a root of a cubic equation with rational coefficients but having no rational root.*

This theorem shows also that it is not possible to construct a regular polygon of 9 sides and hence that it is not possible to construct the angle $40°$ by ruler and compasses. Indeed, if $r = \cos 40° + i \sin 40°$, then $r + 1/r = 2 \cos 40°$ is a root (Ex. 7, p. 83) of

$$z^3 - 3z + 1 = 0.$$

The same equation follows also from the identity

$$\cos 3A = 4 \cos^3 A - 3 \cos A$$

by taking $A = 40°$, replacing $\cos 120°$ by its value $-\frac{1}{2}$, and setting $z = 2 \cos 40°$. Since neither divisor 1 nor -1 of the constant term is a root of the z-cubic, there is no rational root.

Corollary. *It is not possible to trisect every angle by ruler and compasses.* Indeed, angle $40°$ cannot be constructed, while angle $120°$ can be.

9. Duplication of a Cube. Another famous problem of antiquity was the construction of a cube whose volume shall be double that of a given cube. Take the edge of the given cube as the unit of length and denote by x the length of an edge of the required cube. Then $x^3 - 2 = 0$. Since no one of the divisors of 2 is a root of this cubic equation, the theorem stated in § 8 implies the impossibility of the duplication of a cube by ruler and compasses.

10.† Cubic Equations with a Constructible Root. It remains to prove the theorem in § 8 from which we have drawn such important conclusions. Suppose that

$$x^3 + \alpha x^2 + \beta x + \gamma = 0 \qquad (\alpha,\ \beta,\ \gamma \text{ rational}) \tag{20}$$

is a cubic equation having a root x_1 such that a line of length x_1 or $-x_1$ can be constructed by ruler and compasses. We shall prove that one of the roots of (20) is rational.

The construction is in effect the determination of various points as the intersections of auxiliary straight lines and circles. Choose rectangular axes of coördinates. The coördinates of the intersection of two straight lines are rational functions of the coefficients of the equations of the two

lines. To obtain the coördinates of the intersection of the straight line $y = mx + b$ with the circle

$$(x - p)^2 + (y - q)^2 = r^2,$$

we eliminate y and obtain a quadratic equation for x. Thus x, and hence also y, involves no irrationality (besides irrationalities in m, b, p, q, r) other than a square root. Finally, the intersections of two circles are given by the intersections of one of them with their common chord, so that this case reduces to the preceding. Hence the coördinates of the various points located by the construction, and therefore also the length $\pm x_1$ of the segment joining two of them, are found by a finite number of rational operations and extractions of real square roots, performed upon rational numbers and numbers obtained by earlier ones of these operations.

If x_1 is rational, (20) has a rational root as desired. Henceforth, let x_1 be irrational. Then x_1 is the quotient of two sums of terms, each term being a rational number or a rational multiple of a square root. A term may involve superimposed radicals as

$$r = \sqrt{10 - 2\sqrt{5}}, \quad s = \sqrt{10 + 2\sqrt{5}}, \quad t = \sqrt{4 - 2\sqrt{3}}.$$

But t equals $\sqrt{3} - 1$ and would be replaced by that simpler value. As a matter of fact, r is not expressible rationally * in terms of a finite number of square roots of rational numbers, and is said to be a radical of order 2. A term having n superimposed radicals is of *order* n if it is not expressible rationally in terms of radicals each with fewer than n superimposed radicals. In case $x_1 = 2r - 7s$, we would express x_1 in the form $2r - 28\sqrt{5}/r$, involving a single radical of order 2; indeed, $rs = 4\sqrt{5}$. If x_1 involves $\sqrt{3}$, $\sqrt{5}$ and $\sqrt{15}$, we replace $\sqrt{15}$ by $\sqrt{3} \cdot \sqrt{5}$.

We may therefore assume that no one of the radicals of highest order n in x_1 is a rational function with rational coefficients of the remaining radicals of order n and radicals of lower order, that no one of the radicals of order $n - 1$ is a rational function of the remaining radicals of order $n - 1$ and radicals of lower order, etc.

Let $\sqrt{k}$ be a radical of highest order n in x_1. Then

$$x_1 = \frac{a + b\sqrt{k}}{c + d\sqrt{k}},$$

* That is, as a rational integral function with rational coefficients.

where $a, \ldots, d$ do not involve $\sqrt{k}$, but may involve other radicals of order n. If $d \neq 0$, $\sqrt{k} \neq c/d$, in view of the preceding assumption. Thus we may multiply the numerator and denominator of x_1 by $c - d\sqrt{k}$. Hence, whether $d \neq 0$ or $d = 0$, we have

$$x_1 = e + f\sqrt{k} \qquad (f \neq 0),$$

where neither e nor f involves $\sqrt{k}$. Since x_1 is a root of (20), we have $A + B\sqrt{k} = 0$, where A and B are polynomials in $e, f, k, \alpha, \beta, \gamma$. If $B \neq 0$, we could express $\sqrt{k}$ as a rational function $-A/B$ of the remaining radicals in the initial x_1. Hence $B = 0$ and therefore $A = 0$. But the result of substituting $e - f\sqrt{k}$ for x in the cubic function (20) is evidently $A - B\sqrt{k}$. Hence

$$x_2 = e - f\sqrt{k}$$

is a new root of our cubic equation. The third root is

$$x_3 = -\alpha - x_1 - x_2 = -\alpha - 2e.$$

Now α is rational. If e is rational, x_3 is a rational root of (20), as desired. The remaining case is readily excluded. For, if e is irrational, let $\sqrt{s}$ be one of the radicals of highest order in e. Then, as above,

$$x_3 = g + h\sqrt{s} \qquad (h \neq 0),$$

where neither g nor h involves $\sqrt{s}$, while $g - h\sqrt{s}$ is a root $\neq x_3$ of (20), and hence identical with x_1 or x_2. Thus

$$e \pm f\sqrt{k} = g - h\sqrt{s}.$$

Now $\sqrt{s}$ and all the radicals appearing in g, h, s occur in x_3 and hence in e. But $\sqrt{k}$ is not expressible in terms of the remaining radicals of x_1.

We have now proved that if the constructible root x_1 of (20) is irrational, there is a rational root x_3.

11.† Problems such as the trisection of any angle can often be solved by means of certain curves. We note, however, that there exists no plane curve, other than a conic section, whose intersections by an arbitrary straight line can be found by ruler and compasses.*

* J. Petersen, *Algebraische Gleichungen*, p. 169.

CHAPTER IX

Isolation of the Real Roots of an Equation with Real Coefficients

1. Method of Rolle.* *There is at least one real root of* $f'(x) = 0$ *between two consecutive real roots* a *and* b *of* $f(x) = 0$.

For, the graph of $y = f(x)$ has a bend point between a and b.

Corollary. Between two consecutive real roots r and s of $f'(x) = 0$, lies at most one real root of $f(x) = 0$.

For, if there were two such real roots a and b of the latter equation, the first theorem shows that $f'(x) = 0$ would have a real root between a and b and hence between r and s, contrary to hypothesis.

Now $f(x) = 0$ has a real root between r and s if $f(r)$ and $f(s)$ have opposite signs (Ch. I, § 12). Hence the Corollary gives the

Criterion. *If* r *and* s *are consecutive real roots of* $f'(x) = 0$, *then* $f(x) = 0$ *has a single real root between* r *and* s *if and only if* $f(r)$ *and* $f(s)$ *have opposite signs. At most one real root of* $f(x) = 0$ *is greater than the greatest real root of* $f'(x) = 0$, *or less than the least real root of* $f'(x) = 0$.

The final statement follows at once from the first theorem.

Example. For $f(x) = 3x^5 - 25x^3 + 60x - 20$,

$$\tfrac{1}{15} f'(x) = x^4 - 5x^2 + 4 = (x^2 - 1)(x^2 - 4).$$

Hence the roots of $f'(x) = 0$ are ± 1, ± 2. Now

$$f(-\infty) = -\infty,\quad f(-2) = -36,\quad f(-1) = -58,\quad f(1) = 18,\quad f(2) = -4,\quad f(+\infty) = +\infty.$$

Hence there is a single real root in each of the intervals

$$(-1, 1),\quad (1, 2),\quad (2, +\infty),$$

and two imaginary roots. The 3 real roots are positive.

2. The first theorem of § 1 is a special case of

Rolle's Theorem. *Between two consecutive roots* a *and* b *of* $f(x) = 0$, *there is an odd number of real roots of* $f'(x) = 0$, *a root of multiplicity* m *being counted as* m *roots.*

* *Traité de l'algèbre*, Paris, 1690. Hudde knew the method in 1659.

We may argue geometrically, noting that there is an odd number of bend points between a and b, the abscissa of each being a root of $f'(x) = 0$ of odd multiplicity, while the abscissa of an inflexion point with a horizontal tangent is a root of $f'(x) = 0$ of even multiplicity.

To give an algebraic proof, let

$$f(x) \equiv (x-a)^r (x-b)^s Q(x), \qquad a < b,$$

where $Q(x)$ is a polynomial divisible by neither $x - a$ nor $x - b$. Then

$$\frac{(x-a)(x-b)f'(x)}{f(x)} \equiv r(x-b) + s(x-a) + (x-a)(x-b)\frac{Q'(x)}{Q(x)}.$$

The second member has the value $r(a-b) < 0$ for $x = a$ and the value $s(b-a) > 0$ for $x = b$, and hence vanishes an odd number of times between a and b (Ch. I, § 12). But, in the left member, $(x-a)(x-b)$ and $f(x)$ remain of constant sign between a and b, since $f(x) = 0$ has no root between a and b. Hence $f'(x)$ vanishes an odd number of times.

Corollary. If $f(x) = 0$ has only real roots, $f'(x) = 0$ has only real roots distributed as follows: an $(m-1)$-fold root equal to each m-fold root of $f(x) = 0$ for $m \geqq 2$; a single root, which is a simple root, between two consecutive roots of $f(x) = 0$.

For, if the roots of $f(x) = 0$ are $a, b, c, \ldots$, arranged in ascending order, of multiplicities $r, s, t, \ldots$, respectively, then $a, b, c, \ldots$ are roots of $f'(x) = 0$ of multiplicities $r-1, s-1, t-1, \ldots$, and between a and b lies at least one real root of $f'(x) = 0$, etc. The number of these roots of $f'(x) = 0$ is thus at least

$$(r-1) + 1 + (s-1) + 1 + (t-1) + \cdots = r+s+t+ \cdots -1 = n-1,$$

if n is the degree of f. But f' is of degree $n-1$ and hence has only these roots. Thus only one of its roots lies between a and b.

EXERCISES

1. $x^5 - 5x + 2 = 0$ has 1 negative, 2 positive and 2 imaginary roots.
2. $x^6 + x - 1 = 0$ has 1 negative, 1 positive and 4 imaginary roots.
3. $x^5 - 3x^3 + 2x^2 - 5 = 0$ has two imaginary roots, and a real root in each of the intervals $(-2, -1.5)$, $(-1.5, -1)$, $(1, 2)$.
4. $f(x) = 4x^5 - 3x^4 - 2x^2 + 4x - 10 = 0$ has a single real root. Hint:

$$F(x) = \tfrac{1}{4} f'(x) = 5x^4 - 3x^3 - x + 1 = 0$$

has no real root, since $F'(x) = 0$ has a single real root and for it F is positive.

5. If $f^{(k)}(x) = 0$ has imaginary roots, $f(x) = 0$ has imaginary roots.
6. If $f'(x) = 0$ has exactly r real roots, the number of real roots of $f(x) = 0$ is $r + 1$ or is less than $r + 1$ by an even number, a root of multiplicity m being counted as m roots.

3. Sturm's Method. Let $f(x) = 0$ be the given equation with real coefficients, and $f'(x)$ the first derivative of $f(x)$. The first step of the usual process for seeking the greatest common divisor of $f(x)$ and $f'(x)$ consists in dividing f by f' until we obtain a remainder $r(x)$, whose degree is less than that of f'. Then, if q_1 is the quotient, we have $f = q_1 f' + r$. We write $f_2 = -r$, divide f' by f_2, and denote by f_3 the remainder with its sign changed. Thus

$$f = q_1 f' - f_2, \quad f' = q_2 f_2 - f_3, \quad f_2 = q_3 f_3 - f_4, \quad \ldots .$$

The latter equations, in which each remainder is exhibited as the negative of a polynomial f_i, yield a modified process, just as effective as the former process, for finding the greatest common divisor G of $f(x)$ and $f'(x)$ if it exists.

Suppose that $-f_4$ is the first constant remainder. If $f_4 = 0$, then $f_3 = G$, since f_3 divides f_2 and hence also f' and f (by using our equations in reverse order); while conversely, any common divisor of f and f' divides f_2 and hence also f_3.

But if f_4 is a constant $\neq 0$, f and f' have no common divisor involving x. This case arises if and only if $f(x) = 0$ has no multiple root (Ch. I, § 7), and is the only case considered in §§ 4–6.

Before stating Sturm's theorem in general, we shall state it for a numerical case and illustrate its use.

Example. $f(x) = x^3 + 4x^2 - 7$. Then $f' = 3x^2 + 8x$,

$$f = (\tfrac{1}{3}x + \tfrac{4}{9})f' - f_2, \quad f_2 = \tfrac{32}{9}x + 7,$$

$$f' = (\tfrac{27}{32}x + \tfrac{603}{1024})f_2 - f_3, \quad f_3 = \tfrac{4221}{1024}.$$

For $x = 1$, the signs of f, f', f_2, f_3 are $- + + +$, showing a single variation of consecutive signs. For $x = 2$, the signs are $+ + + +$, showing no variation of signs. Sturm's theorem states that there is a *single* real root between 1 and 2. For $x = -\infty$, the signs are $- + - +$, showing 3 variations of signs. The theorem states that there are $3 - 1 = 2$ real roots between $-\infty$ and 1. Similarly,

x	Signs	Variations
-1	$- - + +$	1
-2	$+ - - +$	2
-3	$+ + - +$	2
-4	$- + - +$	3

Hence there is a single real root between -2 and -1, and a single one between -4 and -3. Each real root has now been *isolated* since we have found two numbers such that a single real root lies between these two numbers or equals one of them.

4. Sturm's Theorem. *Let $f(x) = 0$ be an equation with real coefficients and without multiple roots. Modify the usual process for seeking the greatest common divisor of $f(x)$ and its first derivative* $f_1(x)$ by exhibiting each remainder as the negative of a polynomial f_i:*

$$(1) \quad f = q_1f_1 - f_2,\ f_1 = q_2f_2 - f_3,\ f_2 = q_3f_3 - f_4,\ \ldots,\ f_{n-2} = q_{n-1}f_{n-1} - f_n,$$

*where ** f_n is a constant $\neq 0$. If a and b are real numbers, $a < b$, neither a root of $f(x) = 0$, the number of real roots of $f(x) = 0$ between a and b equals the excess of the number of variations of signs of*

$$(2) \qquad f(x),\ f_1(x),\ f_2(x),\ \ldots,\ f_{n-1}(x),\ f_n$$

for $x = a$ over the number of variations of signs for $x = b$. Terms which vanish are to be dropped out before counting the variations of signs.

For brevity, let V_x denote the number of variations of signs of the numbers (2) when x is a particular real number not a root of $f(x) = 0$.

First, if x_1 and x_2 are real numbers such that no one of the continuous functions (2) vanishes for a value of x between x_1 and x_2 or for $x = x_1$ or $x = x_2$, the values of any one of these functions for $x = x_1$ and $x = x_2$ are both positive or both negative (Ch. I, § 12), and therefore $V_{x_1} = V_{x_2}$.

Second, let ρ be a root of $f_i(x) = 0$, where $1 \leqq i < n$. Then

$$(3) \qquad f_{i-1}(x) = q_if_i(x) - f_{i+1}(x)$$

and the equations (1) following this one show that $f_{i-1}(x)$ and $f_i(x)$ have no common divisor involving x (since it would divide the constant f_n). By hypothesis, $f_i(x)$ has the factor $x - \rho$. Hence $f_{i-1}(x)$ does not have this factor $x - \rho$. Thus, by (3),

$$f_{i-1}(\rho) = -f_{i+1}(\rho) \neq 0.$$

Hence, if p is a sufficiently small positive number, the values of

$$f_{i-1}(x), \quad f_i(x), \quad f_{i+1}(x)$$

for $x = \rho - p$ show just one variation of signs, since the first and third values are of opposite signs, and for $x = \rho + p$ show just one variation of

* The notation f_1 instead of the usual f', and similarly f_0 instead of f, is used to regularize the notation of all the f's, and enables us to write any one of the equations (1) in the single notation (3).

** If the division process did not yield ultimately a constant remainder $\neq 0$, f and f_1 would have a common factor involving x, and hence $f(x) = 0$ a multiple root.

signs, and therefore show no change in the number of variations of sign for the two values of x.

It follows from the first and second cases that $V_\alpha = V_\beta$ if α and β are real numbers for neither of which any one of the functions (2) vanishes and such that no root of $f(x) = 0$ lies between α and β.

Third, let r be a root of $f(x) = 0$. By Taylor's Theorem (8) of Ch. I,

$$f(r - p) = -pf'(r) + \tfrac{1}{2}p^2f''(r) - \dots,$$
$$f(r + p) = \quad pf'(r) + \tfrac{1}{2}p^2f''(r) + \dots.$$

If p is a sufficiently small positive number, each of these polynomials in p has the same sign as its first term. For, after removing the factor p, we obtain a quotient of the form $a_0 + s$, where $s = a_1p + a_2p^2 + \dots$ is numerically less than a_0 for all values of p sufficiently small (Ch. I, end of § 11). Hence if $f'(r)$ is positive, $f(r - p)$ is negative and $f(r + p)$ positive, so that the terms $f(x)$, $f_1(x) \equiv f'(x)$ have the signs $- +$ for $x = r - p$ and the signs $+ +$ for $x = r + p$. If $f'(r)$ is negative, these signs are $+ -$ and $- -$ respectively. In each case, $f(x)$, $f_1(x)$ show one more variation of signs for $x = r - p$ than for $x = r + p$. Evidently p may be chosen so small that no one of the functions $f_1(x), \dots, f_n$ vanishes for either $x = r - p$ or $x = r + p$, and such that $f_1(x)$ does not vanish for a value of x between $r - p$ and $r + p$, so that $f(x) = 0$ has the single real root r between these limits (§ 1). Hence by the first and second* cases, $f_1, \dots, f_n$ show the same number of variations of signs for $x = r - p$ and $x = r + p$. Thus, for the entire series of functions (2), we have

$$V_{r-p} - V_{r+p} = 1. \tag{4}$$

The real roots of $f(x) = 0$ within the main interval from a to b (*i.e.*, the aggregate of numbers between a and b) separate it into intervals. By the earlier result, V_x has the same value for all numbers in the same interval. By the present result (4), the value of V_x in any interval ex-

* The argument in the second case when applied for $i = 1$ requires the use of $f_0 = f$ and hence does not indicate the variations in a series lacking f. To avoid the necessity of treating this case $i = 1$, we restricted p further than done at the outset so that $f_1(x)$ shall not vanish between $r - p$ and $r + p$. This necessary step in the proof is usually overlooked. Moreover, we have not adopted the usual argument based upon the continuous change of x from a to b, in view of the ambiguity of V_x when x is a root of $f(x) = 0$, etc.

ceeds the value for the next interval by unity. Hence V_a exceeds V_b by the number of real roots between a and b.

COROLLARY. If $a < b$, $V_a \geqq V_b$.

EXERCISES

Isolate by Sturm's theorem the real roots of

1. $x^3 + 2x + 20 = 0$.   2. $x^3 + x - 3 = 0$.

5. Simplifications of Sturm's Functions. In order to avoid fractions, we may first multiply $f(x)$ by a *positive* constant before dividing it by $f_1(x)$, and similarly multiply f_1 by a positive constant before dividing it by f_2, etc. Moreover, we may remove from any f_i any factor k_i which is either a positive constant or a polynomial in x positive for* $a \leqq x \leqq b$, before we use that f_i as the next divisor.

To prove that Sturm's theorem remains true when these modified functions $f, F_1, \ldots, F_m$ are employed in place of functions (2), consider the equations replacing (1):

$$f_1 = k_1F_1, \quad c_2f = q_1F_1 - k_2F_2, \quad c_3F_1 = q_2F_2 - k_3F_3,$$
$$c_4F_2 = q_3F_3 - k_4F_4, \quad \ldots, \quad c_mF_{m-2} = q_{m-1}F_{m-1} - k_mF_m,$$

in which $c_2, c_3, \ldots$ are positive constants and F_m is a constant $\neq 0$. A common divisor (involving x) of F_{i-1} and F_i would divide $F_{i-2}, \ldots, F_2, F_1, f, f_1$, whereas $f(x) = 0$ has no multiple roots. Hence if ρ is a root of $F_i(x) = 0$, then $F_{i-1}(\rho) \neq 0$ and

$$c_{i+1}F_{i-1}(\rho) = -k_{i+1}(\rho)\,F_{i+1}(\rho), \quad c_{i+1} > 0, \quad k_{i+1}(\rho) > 0.$$

Thus F_{i-1} and F_{i+1} have opposite signs for $x = \rho$. We proceed as in § 4.

EXAMPLE 1. If $f(x) = x^3 + 6x - 10$, $f_1 = 3(x^2 + 2)$ is always positive. Hence we may employ f and $F_1 = 1$. For $x = -\infty$, there is one variation of signs; for $x = +\infty$, no variation. Hence there is a single real root; it lies between 1 and 2.

EXAMPLE 2. If $f(x) = 2x^4 - 13x^2 - 10x - 19$, we may take

$$f_1 = 4x^3 - 13x - 5.$$

Then

$$2f = xf_1 - f_2, \quad f_2 = 13x^2 + 15x + 38 = 13(x + \tfrac{15}{26})^2 + \tfrac{1751}{52}.$$

* Usually we would require that k_i be positive for all values of x, since we usually wish to employ the limits $-\infty$ and $+\infty$.

Since f_2 is always positive, we need go no further (we may take $F_2 = 1$). For $x = -\infty$, the signs are $+ - +$; for $x = +\infty$, $+ + +$. Hence there are two real roots. The signs for $x = 0$ are $- - +$. Hence one real root is positive and the other negative.

EXERCISES

Isolate by Sturm's theorem the real roots of

1. $x^3 + 3x^2 - 2x - 5 = 0$. 2. $x^4 + 12x^2 + 5x - 9 = 0$.
3. $x^3 - 7x - 7 = 0$. 4. $3x^4 - 6x^2 + 8x - 3 = 0$.
5. $x^6 + 6x^5 - 30x^2 - 12x - 9 = 0$ [stop with f_2].
6. $x^4 - 8x^3 + 25x^2 - 36x + 8 = 0$.
7. For $f = x^3 + px + q$ $(p \neq 0)$, $f_1 = 3x^2 + p$, $f_2 = -2px - 3q$,

$$4p^2 f_1 = (-6px + 9q) f_2 - f_3, \quad f_3 = -4p^3 - 27q^2,$$

so that f_3 is the discriminant Δ (Ch. III, § 3). Let $[p]$ denote the sign of p. Then the signs of f, f_1, f_2, f_3 are

$$\begin{array}{ccccl} - & + & +[p] & [\Delta] & \text{for } x = -\infty, \\ + & + & -[p] & [\Delta] & \text{for } x = +\infty. \end{array}$$

For Δ negative there is a single real root. For Δ positive and therefore p negative, there are 3 distinct real roots. For $\Delta = 0$, f_2 is a divisor of f_1 and f, so that $x = -3q/(2p)$ is a double root.

8. If one of Sturm's functions has p imaginary roots, the initial equation has at least p imaginary roots. (Darboux.)

6. Sturm's Functions for a Quartic Equation. For the reduced quartic equation $f(z) = 0$,

$$(5) \qquad \begin{cases} f = z^4 + qz^2 + rz + s, \\ f_1 = 4z^3 + 2qz + r, \\ f_2 = -2qz^2 - 3rz - 4s. \end{cases}$$

Let $q \neq 0$ and divide $q^2 f_1$ by f_2. The negative of the remainder is

$$(6) \qquad f_3 = Lz - 12rs - rq^2, \quad L \equiv 8qs - 2q^3 - 9r^2.$$

Let $L \neq 0$. Then f_4 is a constant which is zero if and only if $f = 0$ has multiple roots, *i.e.*, if its discriminant Δ is zero. We therefore desire f_4 expressed as a multiple of Δ. By Ch. IV, § 4,

$$(7) \qquad \Delta = -4P^3 - 27Q^2, \quad P = -4s - \frac{q^2}{3}, \quad Q = \tfrac{8}{3}qs - r^2 - \tfrac{2}{27}q^3.$$

We may employ P and Q to eliminate

(8) $$4\,s = -P - \frac{q^2}{3}, \quad r^2 = -Q - \tfrac{2}{3}\,qP - \tfrac{8}{27}\,q^3.$$

We divide $L^2 f_2$ by

(9) $$f_3 = Lz + 3\,rP, \quad L \equiv 9\,Q + 4\,qP.$$

The negative of the remainder is

(10) $$18\,r^2qP^2 - 9\,r^2LP + 4\,sL^2 = q^2\Delta.$$

The left member is easily reduced to $q^2\Delta$. Inserting the values (8) and replacing L^2 by $L(9\,Q + 4\,qP)$, we get

$$-18\,qQP^2 - 12\,q^2P^3 - \tfrac{16}{3}\,q^4\,P^2 + 2\,qP^2L + \tfrac{4}{3}\,q^3PL - 3\,q^2QL.$$

Replacing L by its value (9), we get $q^2\Delta$. Hence we may take

(11) $$f_4 = \Delta.$$

Hence if $qL\Delta \neq 0$, we may take (5), (9), (11) as Sturm's functions.

Denote the sign of q by $[q]$. The signs of Sturm's functions are

$$\begin{array}{cccccl} + & - & -[q] & -[L] & [\Delta] & \text{for } x = -\infty, \\ + & + & -[q] & [L] & [\Delta] & \text{for } x = +\infty. \end{array}$$

First, let $\Delta > 0$. If q is negative and L is positive, there are four real roots. In each of the remaining three cases for q and L, there are two variations of signs in either of the two series and hence no real root.

Next, let $\Delta < 0$. In each of the three cases in which q and L are not both positive, there are three variations of signs in the first series and one variation in the second, and hence just two real roots. If q and L are both positive, the number of variations is 1 in the first series and 3 in the second, so that this case is excluded by the Corollary to Sturm's Theorem. To give a direct proof, note that by the value of L in (6), $4\,s > q^2$, and that P is negative by (7), so that each term of (10) is $\geqq 0$, whence $\Delta > 0$.

Hence, if $qL\Delta \neq 0$, there are four distinct real roots if and only if Δ and L are positive, and q negative; two distinct real and two imaginary roots if and only if Δ is negative. See Ex. 5 below.

EXERCISES

1. If $q\Delta \neq 0$, $L = 0$, then $f_3 = 3\,rP$ is not zero and its sign is immaterial in determining the number of real roots: two if $q < 0$, none if $q > 0$. By (10), q has the same sign as Δ.

2. If $r\Delta \neq 0$, $q = 0$, obtain $-f_3$ by substituting $z = -4s/(3r)$ in f_1. Show that we may take $f_3 = r\Delta$ and that there are just two real roots if $\Delta < 0$, no real root if $\Delta > 0$.

3. If $\Delta \neq 0$, $q = r = 0$, there are just two real roots if $\Delta < 0$, no real root if $\Delta > 0$. Since $\Delta = 256\, s^3$, check by solving $z^4 + s = 0$.

4. If $\Delta \neq 0$, $qL = 0$, there are just two real roots if $\Delta < 0$, no real root if $\Delta > 0$. [Combine the results in Exs. 1–3.]

5. If $\Delta < 0$, there are just two real (distinct) roots; if $\Delta > 0$, $q < 0$, $L > 0$, four distinct real roots; if $\Delta > 0$ and either $q \geqq 0$ or $L \leqq 0$, no real root. [Combine the theorem in the text with that in Ex. 4.]

6. Apply the criterion in Ex. 5 to Exs. 2, 4, 6, p. 99.

7. Apply to Exs. 1–3, p. 39, and Exs. 1–4, p. 43.

8. Show that the criterion of Ex. 5 is equivalent to the theorem in Ch. IV, § 7. If $\Delta > 0$, $L > 0$, $q < 0$, then $4s - q^2 < 0$ by (6). Conversely, if $\Delta > 0$, $q < 0$, $4s - q^2 < 0$, then $L > 0$. For, if $L \leqq 0$, $9Q \leqq -4qP < 0$, since $P < 0$ by the value (7) of Δ. Thus $81\,Q^2 \geqq 16\,q^2P^2$, $\Delta \leqq \delta$, where

$$\delta = -4P^3 - \tfrac{16}{3}q^2P^2 = 4P^2(-P - \tfrac{4}{3}q^2) = 4P^2(4s - q^2) < 0,$$

$-P$ having been replaced by its value in (7). Thus $\Delta < 0$, contrary to hypothesis. The two criteria for four real roots are therefore equivalent. The criterion for 2 distinct real and 2 imaginary roots is $\Delta < 0$ in each theorem. By formal logic the criteria for no real root must be equivalent.

9. If α, β, γ are the roots of a cubic equation $f(x) = 0$, Sturm's functions* f, f_1, f_2, f_3 equal products of positive constants by

$$(x-\alpha)(x-\beta)(x-\gamma),\quad \Sigma(x-\beta)(x-\gamma),\quad \Sigma(\alpha-\beta)^2(x-\gamma),\quad (\alpha-\beta)^2(\alpha-\gamma)^2(\beta-\gamma)^2.$$

Why is it sufficient to prove this for a reduced cubic equation?

Take f as in Ex. 7, p. 99. Proof is needed only for the third function. In it the coefficient of x equals $2\,\Sigma\alpha^2 - 2\,\Sigma\alpha\beta = -6\,p$, while the constant is

$$-\Sigma\alpha^2\gamma + 6\,\alpha\beta\gamma = -3\,q - 6\,q,$$

by Ex. 1, p. 64. Thus the third function equals $3\,f_2$.

10. Sturm's functions for any equation with the n roots $\alpha, \beta, \ldots, \pi, \omega$ equal products of positive constants by

$$(x-\alpha)\ldots(x-\omega),\quad \Sigma(x-\beta)\ldots(x-\omega),\quad \Sigma(\alpha-\beta)^2(x-\gamma)\ldots(x-\omega),$$
$$\Sigma(\alpha-\beta)^2(\alpha-\gamma)^2(\beta-\gamma)^2(x-\delta)\ldots(x-\omega),\quad \ldots,\quad (\alpha-\beta)^2\ldots(\pi-\omega)^2.$$

Verify this for $n = 4$, using § 6. A convenient reference to a proof for any n is Salmon's *Modern Higher Algebra*, pp. 49–53.

11. There are as many pairs p of imaginary roots as there are variations of signs in the leading coefficients of Sturm's functions, *i.e.*, $p = V_{+\infty}$. Hints: Of any two consecutive Sturm's functions, the one of even degree has the same signs for $x = -\infty$ and $x = +\infty$, while the one of odd degree has opposite signs.

* In Exs. 9–12, it is assumed that there are $n + 1$ Sturm's functions for the equation of degree n.

Hence, for the two functions, $V_{-\infty} + V_{+\infty} = 1$. There are n pairs of consecutive Sturm's functions.

Hence $V_{-\infty} + V_{+\infty} = n$, the degree of the equation.

Subtract $V_{-\infty} - V_{+\infty} = r$, the number of real roots.

Thus $2\, V_{+\infty} = n - r = 2\, p$.

12. By Exs. 10, 11, the number of pairs of imaginary roots is the number of variations of signs in the series

$$1,\ n,\ \Sigma(\alpha-\beta)^2,\quad \Sigma(\alpha-\beta)^2(\alpha-\gamma)^2(\beta-\gamma)^2,\ \ldots,$$

provided no one of these sums is zero.

7.† Sturm's Theorem for the Case of Multiple Roots. Let* $f_n(x)$ be the greatest common divisor of $f(x)$ and $f_1 = f'(x)$. We have equations (1) in which f_n is now not a constant. *The difference $V_a - V_b$ is the number of real roots between a and b, each multiple root being counted only once.*

If ρ is a root of $f_i(x) = 0$, but not a multiple root of $f(x) = 0$, then $f_{i-1}(\rho) \neq 0$. For, if it were zero, $x - \rho$ would by (1) be a common factor of f and f_1. We may now proceed as in the second case in § 4.

The third case requires a modified proof only when r is a multiple root. Let r be a root of multiplicity m, $m \geqq 2$. Then $f(r), f'(r), \ldots, f^{(m-1)}(r)$ are zero and, by Taylor's Theorem,

$$f(r+p) = \frac{p^m}{1 \cdot 2 \cdots m} f^{(m)}(r) + \cdots,$$

$$f'(r+p) = \frac{p^{m-1}}{1 \cdot 2 \cdots (m-1)} f^{(m)}(r) + \cdots.$$

These have like signs if p is a positive number so small that the signs of the polynomials are those of their first terms. Similarly, $f(r-p)$ and $f'(r-p)$ have opposite signs. Hence f and f_1 show one more variation of signs for $x = r - p$ than for $x = r + p$. Now $(x-r)^{m-1}$ is a factor of f and f_1 and hence, by (1), of $f_2, \ldots, f_n$. Let their quotients by this factor be $\phi, \phi_1, \ldots, \phi_n$. Then equations (1) hold after the f's are replaced by the ϕ's. Taking p so small that $\phi_1(x) = 0$ has no root between $r - p$ and $r + p$, we see by the first and second cases in § 4 that $\phi_1, \ldots, \phi_n$ show the same number of variations of signs for $x = r - p$ as for $x = r + p$. The same is true for $f_1, \ldots, f_n$ since the products of $\phi_1, \ldots, \phi_n$ by $(x-r)^{m-1}$ have for a given x the same signs as $\phi_1 \ldots, \phi_n$ or the same signs as $-\phi_1, \ldots, -\phi_n$. But the latter series evidently shows the same number of variations of signs as $\phi_1, \ldots, \phi_n$. Hence (4) is proved and consequently the present theorem.

* The degree of $f(x)$ is not n, nor was it necessarily n in § 4.

EXERCISES

1.† For $f = x^4 - 8x^2 + 16$, $f_1 = x^3 - 4x$, $f_2 = x^2 - 4$, $f_1 = xf_2$. Hence $n = 2$. Then $V_{-\infty} = 2$, $V_{\infty} = 0$, and there are only two real roots, each a double root.

2.† $f = (x-1)^3(x-2)$. 3.† $(x-1)^2(x+2)^3$. 4.† $x^4 - x^2 - 2x + 2$.

8.† Budan's Theorem. *Let a and b be real numbers, $a < b$, neither a root of $f(x) = 0$, an equation of degree n with real coefficients. Let V_a denote the number of variations of signs of*

$$(12) \qquad f(x), \quad f'(x), \quad f''(x), \quad \ldots, \quad f^{(n)}(x)$$

for $x = a$, after vanishing terms have been deleted. Then $V_a - V_b$ is either the number of real roots of $f(x) = 0$ between a and b or exceeds the number of those roots by an even integer. A root of multiplicity m is here counted as m roots.

In case $V_a - V_b$ is 0 or 1, it is the exact number of real roots between a and b. In other cases, it is merely an upper limit to the number of those roots. While therefore the present method is not certain to lead to the isolation of the real roots, it is simpler to apply than Sturm's method. Indeed, for an equation of degree 6 or 7 with simple coefficients, Sturm's functions may introduce numbers of 50 or more figures.

The proof is quite simple if no term of the series (12) vanishes for $x = a$ or for $x = b$ and if no two consecutive terms vanish for the same value of x between a and b. Indeed, if no one of the terms vanishes for $x_1 \leqq x \leqq x_2$, then $V_{x_1} = V_{x_2}$, since any term has the same sign for $x = x_1$ as for $x = x_2$. Next, let r be a root of $f^{(i)}(x) = 0$, $a < r < b$. By hypothesis, the first derivative $f^{(i+1)}(x)$ of $f^{(i)}(x)$ is not zero for $x = r$. As in the third step (now actually the case $i = 0$) in § 4, $f^{(i)}(x)$ and $f^{(i+1)}(x)$ show one more variation of signs for $x = r - p$ than for $x = r + p$, where p is a sufficiently small positive number. If $i > 1$, $f^{(i)}$ is preceded by a term $f^{(i-1)}$ in (12). By hypothesis, $f^{(i-1)}(x) \neq 0$ for $x = r$ and hence has the same sign for $x = r - p$ and $x = r + p$ when p is sufficiently small. For these values of x, $f^{(i)}(x)$ has opposite signs. Hence $f^{(i-1)}$ and $f^{(i)}$ show one more or one less variation of signs for $x = r - p$ than for $x = r + p$, so that $f^{(i-1)}$, $f^{(i)}$, $f^{(i+1)}$ show two more variations or the same number of variations of signs.

Next, let no term of the series (12) vanish for $x = a$ or for $x = b$, but let several successive terms

$$(13) \qquad f^{(i)}(x), \quad f^{(i+1)}(x), \quad \ldots, \quad f^{(i+j-1)}(x)$$

all vanish for a value r of x between a and b, while $f^{(i+j)}(r)$ is not zero, say positive.* Let I_1 be the interval between $r - p$ and r, and I_2 the interval between r and $r + p$. Let the positive number p be so small that no one of the functions (13) or $f^{(i+j)}(x)$ is zero in these intervals, so that the last function remains positive. Hence $f^{(i+j-1)}(x)$ increases with x (since its derivative is positive) and is therefore negative in I_1 and positive in I_2. Thus $f^{(i+j-2)}(x)$ decreases in I_1 and increases in I_2 and hence is positive in each interval. In this manner we may verify the signs in the following table:

	$f^{(i)}$	$f^{(i+1)}$	$f^{(i+2)}$	$\cdots$	$f^{(i+j-3)}$	$f^{(i+j-2)}$	$f^{(i+j-1)}$	$f^{(i+j)}$
I_1	$(-1)^j$	$(-1)^{j-1}$	$(-1)^{j-2}$	$\cdots$	$-$	$+$	$-$	$+$
I_2	$+$	$+$	$+$	$\cdots$	$+$	$+$	$+$	$+$

Hence these functions show j variations of signs in I_1 and none in I_2.

If $i > 0$, the first term of (13) is preceded by a function $f^{(i-1)}(x)$ which is not zero for $x = r$, and hence not zero in I_1 or I_2 if p is sufficiently small. If j is even, the signs of $f^{(i-1)}$ and $f^{(i)}$ are $+\,+$ or $-\,+$ in both I_1 and I_2, showing no loss in the number of variations of signs. If j is odd, their signs are

$$\begin{array}{c|c} I_1 & +\,- \\ I_2 & +\,+ \end{array} \quad \text{or} \quad \begin{array}{c} -\,- \\ -\,+ \end{array}$$

so that there is a gain or loss of a single variation of signs. Hence

$$f^{(i-1)},\quad f^{(i)},\quad f^{(i+1)},\ \ldots,\ f^{(i+j)}$$

show a loss of j variations of signs if j is even, and a loss of $j \pm 1$ if j is odd, and hence always a loss of an even number $\geqq 0$ of variations of signs.

If $i = 0$, $f^{(i)} \equiv f$ has r as a j-fold root and the functions in the table show j more variations of signs for $x = r - p$ than for $x = r + p$.

Thus, when no one of the functions (12) vanishes for $x = a$ or for $x = b$, the theorem follows as at the end of § 4 (with unity replaced by the multiplicity of a root).

Finally, let one of the functions (12), other than $f(x)$ itself, vanish for $x = a$ or for $x = b$. If δ is a sufficiently small positive number, all of the N roots of $f(x) = 0$ between a and b lie between $a + \delta$ and $b - \delta$, and

* If negative, all signs in the table below are to be changed; but the conclusion holds.

for the latter values no one of the functions (12) is zero. By the above proof,

$$V_{a+\delta} - V_{b-\delta} = N + 2t,$$

$$V_a - V_{a+\delta} = 2j, \quad V_{b-\delta} - V_b = 2s,$$

where t, j, s are integers $\geqq 0$. Hence $V_a - V_b = N + 2(t + j + s)$.

Example. For $f = x^3 - 7x - 7$,

$$f' = 3x^2 - 7, \quad f'' = 6x, \quad f''' = 6.$$

There is one variation of signs for $x = 3$, but none for $x = 4$, so that just one real root lies between 3 and 4. For

x	f	f'	f''	f'''	
-2	-1	$+5$	-12	$+6$	3 variations
-1	-1	-4	-6	$+6$	1 variation.

Thus there are two real roots or no real root between -2 and -1. The former is the case. The reader should isolate the two roots by finding an intermediate value of x for which the series shows two variations of signs.

EXERCISES

Isolate by Budan's theorem the real roots of

1.† $x^3 - x^2 - 2x + 1 = 0$. 2.† $x^3 + 3x^2 - 2x - 5 = 0$.

3.† If $f(a) \neq 0$, V_a equals the number of real roots $> a$ or exceeds that number by an even integer.

4.† There is no root greater than a number making each of the functions (12) positive, if the leading coefficient of $f(x)$ is positive. (Newton.)

5.† Divide $f(x) = x^n + a_1x^{n-1} + \cdots$ by $x - \alpha$; then

$$f(x) \equiv (x - \alpha)\{x^{n-1} + x^{n-2}g_1(\alpha) + \cdots + g_{n-1}(\alpha)\} + f(\alpha),$$

where $g_1(\alpha) = \alpha + a_1$, $g_2(\alpha) = \alpha^2 + a_1\alpha + a_2, \ldots$. If α is chosen so that $g_1(\alpha), \ldots, g_{n-1}(\alpha), f(\alpha)$ are all positive, no positive root of $f(x) = 0$ exceeds α. (Laguerre.)

9. Descartes' Rule of Signs. *The number of positive roots of an equation with real coefficients either equals the number V of variations of signs in the series of coefficients or is less than V by an even integer. A root of multiplicity m is here counted as m roots.*

For example, $x^6 - 3x^2 + x + 1 = 0$ has either two or no positive roots, the exact number not being found. But $-3x^3 + x + 1 = 0$ has exactly one positive root.

Consider any equation with real coefficients

$$f(x) \equiv a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = 0,$$

with $a_n \neq 0$. For $x = 0$ the functions (12) have the same signs as

$$a_n, \; a_{n-1}, \; \ldots, \; a_1, \; a_0,$$

so that $V_0 = V$. For $x = +\infty$, the functions have the same sign (that of a_0). Thus $V_0 - V_\infty = V$ is either the number of positive roots or exceeds that number by an even integer. Next, the theorem holds if $f(0) = 0$, as shown by removing the factors x.

Corollary. The number of negative roots of $f(x) = 0$ is either the number of variations of signs in the coefficients of $f(-x)$ or is less than that number by an even integer.

Thus $x^6 - 3x^2 + x + 1 = 0$ has either two or no negative roots, since $x^6 - 3x^2 - x + 1 = 0$ has two or no positive roots.

EXERCISES

1. $x^3 - 3x + 2 = 0$ has one negative root and two equal positive roots.
2. $x^3 + a^2x + b^2 = 0$ has two imaginary roots if $b \neq 0$.
3. For n even, $x^n - 1 = 0$ has only two real roots.
4. For n odd, $x^n - 1 = 0$ has only one real root.
5. For n even, $x^n + 1 = 0$ has no real root; for n odd, only one.
6. $x^4 + 12x^2 + 5x - 9 = 0$ has just two imaginary roots.
7. $x^4 + a^2x^2 + b^2x - c^2 = 0$ $(c \neq 0)$ has just two imaginary roots.
8. To find an upper limit to the number of real roots of $f(x) = 0$ between a and b, set

$$x = \frac{a + by}{1 + y} \qquad \left(\therefore y = \frac{x - a}{b - x}\right),$$

multiply by $(1 + y)^n$, and apply Descartes' Rule to the resulting equation in y.

10.† Fourier's Method. If Budan's Theorem gives a loss of two or more variations of signs in passing from a to a larger value b, and hence leaves in doubt the number of real roots between a and b, we may employ a supplementary discussion.

First, let f, f', f'' show two variations of signs at a and no variation at b, while the series beginning with f'' shows no loss in variations (as in the Example in § 8). Then f'' is of constant sign between a and b, and the

graph of $y = f(x)$ has a (single) maximum or minimum point between a and b, according as f'' is negative or positive. If the sum

$$\frac{f(b)}{f'(b)} - \frac{f(a)}{f'(a)}$$

of the subtangents at the points with the abscissas a and b is $> b - a$, the tangents cross before meeting the x-axis and the graph does not intersect the x-axis between a and b, so that there are two imaginary roots in view of Budan's Theorem and

$$(14) \qquad n = V_{-\infty} - V_{\infty} = (V_{-\infty} - V_a) + (V_a - V_b) + (V_b - V_{\infty}).$$

In the contrary case, we examine the value half way between a and b, etc. Clearly the case of imaginary roots will disclose itself after a very few such steps.

Next, in the general case, we shall find, after a suitable subdivision of the interval, three consecutive functions

$$f^{(j)}, \quad f^{(j+1)}, \quad f^{(j+2)}$$

showing two variations of signs at a' and no variation at b', while the later terms of the series show no loss in variations of signs. We may therefore decide as in the first case whether there are two real roots of $f^{(j)} = 0$ in the interval $[a', b']$ or not, and in the latter alternative conclude that $f = 0$ has two imaginary roots.*

EXAMPLE. Let $f(x) = x^5 - 5x^4 - 16x^3 + 12x^2 - 9x - 5$. Then

$$\begin{aligned} f'(x) &= 5x^4 - 20x^3 - 48x^2 + 24x - 9, \\ \tfrac{1}{4}f''(x) &= 5x^3 - 15x^2 - 24x + 6, \\ \tfrac{1}{12}f'''(x) &= 5x^2 - 10x - 8, \\ \tfrac{1}{120}f''''(x) &= x - 1, \quad f^{(5)}(x) = 120. \end{aligned}$$

There is just one real root in each of the intervals $(-3, -2)$, $(-1, 0)$, $(7, 8)$. The interval $(0, 1)$ is in doubt, the signs being

$$\begin{array}{cccccccc} - & - & + & - & - & + & \text{for } x = 0, \\ - & - & - & - & - & + & \text{for } x = 1, \end{array}$$

where 0 is read $-$. The j of the text is here 1. Now

$$\frac{f'(1)}{f''(1)} - \frac{f'(0)}{f''(0)} = \frac{-48}{4(-28)} + \frac{9}{4(6)} = \frac{3}{7} + \frac{3}{8} < 1,$$

* For further details, see Serret, *Algèbre Supérieure*, ed. 4, I, pp. 305–318.

so that we must subdivide the interval. For $x = \frac{1}{2}$, the signs are the same as for $x = 1$. Thus the loss in variations of signs occurs in the interval $(0, \frac{1}{2})$. Now

$$\frac{f'(\frac{1}{2})}{f''(\frac{1}{2})} - \frac{f'(0)}{f''(0)} = \frac{-11\frac{3}{16}}{4\,(-9\frac{1}{8})} + \frac{3}{8} > \frac{1}{2}.$$

Hence there are two imaginary roots.

EXERCISES

1.† $x^5 - 3x^4 + 2x^3 - 8x^2 + 3x - 25 = 0$ has 4 imaginary roots.

2.† $x^6 + x^5 - x^4 - x^3 + x^2 - x + 1 = 0$ has 6 imaginary roots.

CHAPTER X

Solution of Numerical Equations

1. Newton's Method. To find the root between 2 and 3 of

$$x^3 - 2x - 5 = 0,$$

Newton * replaced x by $2 + p$ and obtained

$$p^3 + 6p^2 + 10p - 1 = 0.$$

Since p is a decimal, he neglected** the first two terms and set $10p - 1 = 0$, so that $p = 0.1$, approximately. Replacing p by $0.1 + q$ in the preceding cubic equation, he obtained

$$q^3 + 6.3q^2 + 11.23q + 0.061 = 0.$$

Dividing -0.061 by 11.23, he obtained -0.0054 as the approximate value of q. Neglecting q^3 and replacing q by $-0.0054 + r$, he obtained

$$6.3r^2 + 11.16196r + 0.000541708 = 0.$$

Dropping $6.3r^2$, he found r and hence

$$x = 2 + 0.1 - 0.0054 - 0.00004853 = 2.09455147.$$

This value is in fact correct to the seventh decimal place. But the method will not often lead as quickly to so accurate a value of the root.

The method is usually presented in the following form. Given that a is an approximate value of a real root of $f(x) = 0$, we can usually find a nearer approximation $a + h$ to the root by neglecting the powers $h^2, h^3, \ldots$ of the small number h in Taylor's formula

$$f(a + h) = f(a) + f'(a)h + f''(a)\frac{h^2}{2} + \cdots$$

and hence by taking

$$f(a) + f'(a)h = 0, \quad h = \frac{-f(a)}{f'(a)}.$$

We then repeat the process with $a + h$ in place of the former a.

* Isaaci Newtoni, *Opuscula*, I, 1794, p. 10, p. 37 [found before 1676].

** At this early stage of the work it is usually safer to retain also the term in p^2 and thus find p approximately by solving a quadratic equation.

Thus in Newton's example, we have, for $a = 2$,

$$h = \frac{-f(2)}{f'(2)} = \frac{1}{10}, \quad a' = a + h = 2.1,$$

$$h' = \frac{-f(2.1)}{f'(2.1)} = \frac{-0.061}{11.23} = -0.0054 \ldots .$$

2. Graphical Discussion of Newton's Method. Using rectangular coördinates, consider the graph of $y = f(x)$ and the point P on it with the abscissa $OQ = a$ (Fig. 22). Let the tangent at P meet the x-axis at T

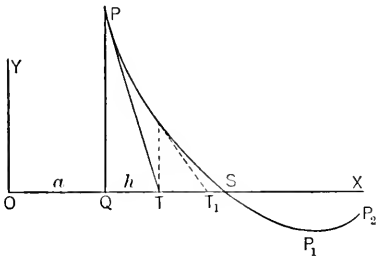


Fig. 22

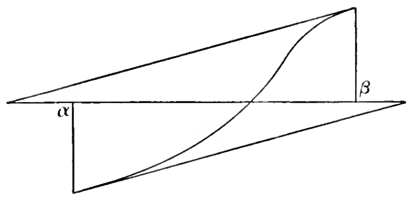


Fig. 23

and let the graph meet the x-axis at S. Take $h = QT$, the subtangent. Then

$$QP = f(a), \quad f'(a) = \tan XTP = -f(a)/h,$$

$$h = \frac{-f(a)}{f'(a)}.$$

In the fictitious graph in Fig. 22, $OT = a + h$ is a better approximation to the root OS than $OQ = a$. The next step (indicated by dotted lines) gives a still better approximation OT_1.

If, however, we had begun with the abscissa a of a point P_1 near a bend point, the subtangent would be very large and the method would probably fail to give a better approximation. Failure is certain if we use a point P_2 such that a single bend point lies between it and S.

We are concerned with the approximation to a root previously isolated as the only real root between two given numbers α and β. These should be chosen so nearly equal that $f'(x) = 0$ has no real root between α and β, and hence $f(x) = 0$ no bend point between α and β. Further, if $f''(x) = 0$

has a root between our limits, our graph will have an inflexion point with an abscissa between α and β, and the method likely will fail (Fig. 23).

Let, therefore, neither $f'(x)$ nor $f''(x)$ vanish between α and β. Since f'' preserves its sign in the interval from α to β, while f changes in sign, f'' and f will have the same sign for one end point. According as the abscissa of this point is α or β, we take $a = \alpha$ or $a = \beta$ for the first step of Newton's process. In fact, the tangent at one of the end points meets the x-axis at a point T with an abscissa within the interval from α to β. If $f'(x)$ is positive in the interval, we have Fig. 24 or Fig. 25; if f' is negative, Fig. 26 or Fig. 22.

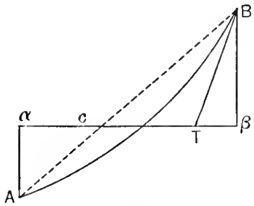


Fig. 24

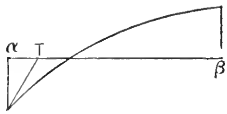


Fig. 25

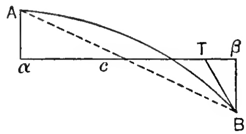


Fig. 26

In Newton's example, the graph between the points with the abscissas $\alpha = 2$ and $\beta = 3$ is of the type in Fig. 24, but more nearly like a vertical straight line. In view of this feature of the graph, we may safely take $a = \alpha$, as did Newton, although our general procedure would be to take $a = \beta$. The next step, however, accords with our present process; we have $\alpha = 2$, $\beta = 2.1$ in Fig. 24 and hence we now take $a = \beta$, getting

$$\frac{0.061}{11.23} = 0.0054$$

as the subtangent, and hence $2.1 - 0.0054$ as the approximate root.

If we have secured (as in Fig. 24 or Fig. 26) a better upper limit to the root than β, we may take the abscissa c of the intersection of the chord AB with the x-axis as a better lower limit than α. By similar triangles,

$$-f(\alpha) : c - \alpha = f(\beta) : \beta - c,$$

$$c = \frac{\alpha f(\beta) - \beta f(\alpha)}{f(\beta) - f(\alpha)}. \tag{1}$$

This method of finding the value of c intermediate to α and β is called the method of interpolation (regula falsi).

In Newton's example, $\alpha = 2$, $\beta = 2.1$,

$$f(\alpha) = -1, \quad f(\beta) = 0.061, \quad c = 2.0942.$$

The advantage of having c at each step is that we know a close limit of the error made in the approximation to the root.

We may combine the various possible cases discussed into one:

If $f(x) = 0$ has a single real root and $f'(x) = 0$, $f''(x) = 0$ have no real root between α and β, and if we designate by β that one of the numbers α and β for which $f(\beta)$ and $f''(\beta)$ have the same sign, then the root lies in the narrower interval from c to $\beta - f(\beta)/f'(\beta)$, where c is given by (1).

It is possible to prove* this theorem algebraically and to show that by repeated applications of it we can obtain two limits α', β' between which the root lies, such that $\alpha' - \beta'$ is numerically less than any assigned positive number. Hence the root can be found in this manner to any desired accuracy.

EXAMPLE. $f(x) = x^3 - 2x^2 - 2$, $\alpha = 2\frac{1}{4}$, $\beta = 2\frac{1}{2}$. Then

$$f(\alpha) = -\tfrac{47}{64}, \quad f(\beta) = \tfrac{9}{8}.$$

Neither of the roots 0, $4/3$ of $f'(x) = 0$ lies between α and β, so that $f(x) = 0$ has a single real root between these limits (Ch. IX, § 1). Nor is the root $\frac{2}{3}$ of $f''(x) = 0$ within these limits. The conditions of the theorem are therefore satisfied. For $\alpha < x < \beta$, the graph is of the type in Fig. 24. We find that

$$c = \tfrac{559}{238} = 2.349, \quad \beta' = \beta - \frac{f(\beta)}{f'(\beta)} = 2.3714,$$

$$\beta' - \frac{f(\beta')}{f'(\beta')} = 2.3597.$$

For $x = 2.3593$, $f(x) = -0.00003$. We therefore have the root to four decimal places. For $a = 2.3593$,

$$f'(a) = 7.2620, \quad a - \frac{f(a)}{f'(a)} = 2.3593041,$$

which is the value of the root correct to 7 decimal places. For, if we change the final digit from 1 to 2, the result is greater than the root in view of our work, while if we change it to 0, $f(x)$ is negative.

* Weber's *Algebra*, 2d ed., I, pp. 380–382; *Kleines Lehrbuch der Algebra*, 1912, p. 163.

EXERCISES

(Preserve the numerical work for later use.)

1. Find the root between 1 and 2 of $x^3 + 4x^2 - 7 = 0$ correct to 7 decimal places.
2. Find the root between -1 and -2 to 5 decimal places.
3. Find a root of $x^3 + 2x + 20 = 0$ to 5 decimal places.

3. Systematic Computation by Newton's Method. Set

$$f_2 = \tfrac{1}{2}f'', \quad f_3 = \frac{1}{2\cdot 3}f''' = \tfrac{1}{3}f_2', \quad f_4 = \frac{1}{2\cdot 3\cdot 4}f'''' = \tfrac{1}{4}f_3', \quad \ldots .$$

Then, by Taylor's formula,

$$\begin{aligned}
f(x+h) &= f(x) + hf'(x) + h^2f_2(x) + h^3f_3(x) + h^4f_4(x) + \cdots .\\
f'(x+h) &= f'(x) + 2hf_2(x) + 3h^2f_3(x) + 4h^3f_4(x) + \cdots .\\
f_2(x+h) &= f_2(x) + 3hf_3(x) + 6h^2f_4(x) + \cdots .\\
f_3(x+h) &= f_3(x) + 4hf_4(x) + \cdots .
\end{aligned}$$

The second formula may also be derived from the first by differentiation with respect to h (or if we prefer, with respect to x), and likewise the third from the second, with a subsequent division by 2, etc.

The work of finding $f(x+h)$, $f'(x+h)$, ... from $f(x)$, $f'(x)$, $f_2(x)$, ... may be arranged as follows for the case $n = 3$, whence $f_4 = 0$:

$$\begin{array}{ll|c|c}
f_3 & f_2 & f' & f \\
 & +hf_3 & +h(f_2+hf_3) & +h(f'+hf_2+h^2f_3) \\ \hline
 & f_2+hf_3 & f'+hf_2+h^2f_3 & f+hf'+h^2f_2+h^3f_3 \\
 & +hf_3 & +h(f_2+2hf_3) & =f(x+h) \\ \hline
 & f_2+2hf_3 & f'+2hf_2+3h^2f_3 = f'(x+h) & \\
 & +hf_3 & & \\ \hline
 & f_2+3hf_3 = f_2(x+h) & &
\end{array}$$

Here we have added hf_3 to f_2. This sum is multiplied by h and the product added to f'. To the resulting sum is added h times the second sum $f_2 + 2hf_3$ in the second column; etc.

Example 1. $f(x) = x^3 - 2x^2 - 2$. Then

$$f'(x) = 3x^2 - 4x, \quad f_2(x) = 3x - 2, \quad f_3(x) = 1.$$

Their values for $x = \beta = 2\frac{1}{2}$ are given in the first line below. Since* $h = -f/f' = -0.129$, the work is as follows:

* Ordinarily we would use at this step the value $h = -.13$, which is sufficiently exact and simplifies the numerical work.

1	5.5	8.75	1.125
	−0.129	−0.69286	−1.03937
	5.371	8.05714	0.08563
	−0.129	−0.67622	
	5.242	7.38092	
	−0.129		
1	5.113		

The numbers at the bottom are the values of

$$f_3, \quad f_2(\beta'), \quad f'(\beta'), \quad f(\beta') \quad \text{for } \beta' = \beta + h = 2.371.$$

Example 2. Netto treats in his *Algebra* the equation

$$f(x) = x^4 + x^3 - 3x^2 - x - 4 = 0.$$

Then

$$f'(x) = 4x^3 + 3x^2 - 6x - 1, \quad f_2 = 6x^2 + 3x - 3, \quad f_3 = 4x + 1, \quad f_4 = 1.$$

Since $f(1) = -6$, $f(2) = 6$, there is a root of $f(x) = 0$ between 1 and 2. By Descartes' Rule, $f'(x)$ and $f_2(x)$ each have a single positive root. Since $f'(1) = 0$, $f_2(1) = 6$, $f_2(2) = 27$, neither has a root between 1 and 2. Since $f(2)$ and $f''(2)$ are of like sign, we take $\beta = 2$. The values of $f_4, \ldots, f$ for $x = 2$ are given in the first line below.

1	9	27	31	6	$\frac{-6}{31} = -0.2$
	−0.2	− 1.76	− 5.048	−5.1904	
	8.8	25.24	25.952	0.8096	
	−0.2	− 1.72	− 4.704		
	8.6	23.52	21.248		
	−0.2	− 1.68			
	8.4	21.84			$\frac{-0.8096}{21.248}$
	−0.2				
1	8.2				$= -0.04$
	−0.04	− 0.3264	− 0.860544	−0.81549824	
	8.16	21.5136	20.387456	−0.00589824	
	−0.04	− 0.3248	− 0.847552		
	8.12	21.1888	19.539904		
	−0.04	− 0.3232			
	8.08	20.8656			
	−0.04		$\frac{0.00589824}{19.539904} = 0.000302-$		
1	8.04				

The root is $2 - 0.2 - 0.04 + 0.000302 = 1.760302$, in which only the last figure is in doubt. Indeed, it can be proved that *if the quotient f/f' begins with k zeros when expressed as a decimal, the best approximation is obtained by carrying the division to $2\,k$ decimal places.*

EXERCISES

1. Extend the work of Example 1 above.
2. Apply the present method to Exs. 1, 2, 3, page 113.
3. Treat in this way Newton's example (§ 1).
4. In the four long formulas at the beginning of § 3, any arithmetical coefficient equals the sum of the one preceding it and the one above that preceding one, as $6 = 3 + 3$, $4 = 1 + 3$.

4. Horner's Method.* To find the root between 2 and 3 of

$$x^3 - 2\,x - 5 = 0$$

by the method now to be explained, we shall modify in two respects the process used by Newton (§ 1). While in the latter process we set $x = 2 + p$ and found the cube of $2 + p$, etc., in order to form the *transformed* equation

$$p^3 + 6\,p^2 + 10\,p - 1 = 0$$

for p, we shall now obtain this equation by a different process. Since $p = x - 2$,

$$x^3 - 2\,x - 5 \equiv (x-2)^3 + 6\,(x-2)^2 + 10\,(x-2) - 1,$$

identically in x. Hence -1 is the remainder obtained when $x^3 - 2\,x - 5$ is divided by $x - 2$; the quotient Q evidently equals

$$(x-2)^2 + 6\,(x-2) + 10.$$

Similarly, 10 is the remainder obtained when this Q is divided by $x - 2$ and the quotient Q_1 equals $(x - 2) + 6$. Another division gives the remainder 6. Hence to find the coefficients 6, 10, -1 of the terms after p^3 in the new equation in the variable $p = x - 2$, we have only to divide the given function $x^3 - 2\,x - 5$ by $x - 2$, the quotient Q by $x - 2$, etc., and take the remainders -1, 10, 6 in reverse order. However, when the work is performed as tabulated below, no reversal of order is needed, since the coefficients then appear on the page in their desired order.

* W. G. Horner, *London Philosophical Transactions*, 1819.

Synthetic Division. We next explain a brief method of performing a division by $x - 2$ and, in general, by $x - h$. When we divide

$$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n$$

by $x - h$, let the constant remainder be r and the quotient be

$$q(x) = b_0x^{n-1} + b_1x^{n-2} + \cdots + b_{n-1}.$$

Comparing the coefficients of $f(x)$ with those in

$$(x - h)\, q(x) + r$$
$$= b_0x^n + (b_1 - hb_0)x^{n-1} + (b_2 - hb_1)\, x^{n-2} + \cdots + (b_{n-1} - hb_{n-2})x + r - hb_{n-1},$$

we obtain relations which may be written in the form

$$b_0 = a_0,\ \ b_1 = a_1 + hb_0,\ \ b_2 = a_2 + hb_1,\ \ \ldots,\ \ b_{n-1} = a_{n-1} + hb_{n-2},\ \ r = a_n + hb_{n-1}.$$

The steps in the work of computing the b's may be tabulated as follows:

$$\begin{array}{cccccc|l}
a_0 & a_1 & a_2 & \ldots & a_{n-1} & a_n & \underline{h} \\
 & hb_0 & hb_1 & \cdots & hb_{n-2} & hb_{n-1} & \\
\hline
b_0 & b_1 & b_2 & \ldots & b_{n-1}, & r &
\end{array}$$

In the second space below a_0 we write b_0 (which equals a_0). Then multiply b_0 by h and enter the product under a_1, add and write the sum b_1 below it, etc. This process was used in Ch. I, § 5, to get the value r of $f(h)$. See also Ch. VI, § 6.

In our example, the work is as follows:

$$\begin{array}{cccc|l}
1 & 0 & -2 & -5 & \underline{2} \\
 & 2 & 4 & 4 & \\
\hline
1 & 2 & 2 & -1 & \\
 & 2 & 8 & & \\
\hline
1 & 4 & 10 & & \\
 & 2 & & & \\
\hline
1 & 6 & & &
\end{array}$$

Thus 1, 6, 10, -1 are the coefficients of the equation in p.

But there is a more essential difference between the methods of Horner and Newton than the detail as to the actual work of finding the transformed equations. Newton used the close approximation 0.1 to the root of the equation in p. As this value exceeds the root p and hence would

lead to a negative correction at the next step, Horner would have used the approximation 0.09 (taking a decimal, with a single significant figure, just less than the root). The next steps of Horner's process are as follows:

$$
\begin{array}{llllll}
1 & 6 & 10 & -1 & \underline{|0.09} & \\
 & 0.09 & 0.5481 & 0.949329 & & \\
\hline
1 & 6.09 & 10.5481 & -0.050671 & & \\
 & 0.09 & 0.5562 & & & \\
\hline
1 & 6.18 & 11.1043 & & & 0.05 \\
 & 0.09 & & & & \overline{11.1} \\
\hline
1 & 6.27 & & & & =0.004 \\
 & 0.004 & 0.025096 & 0.044517584 & & \\
\hline
1 & 6.274 & 11.129396 & -0.006153416 & & \\
 & 0.004 & 0.025112 & & & \\
\hline
1 & 6.278 & 11.154508 & & & \\
 & 0.004 & & & & \\
\hline
1 & 6.282 & & & &
\end{array}
$$

Hence $x=2.094+t$, where t is between 0.0005 and 0.0006. Thus $t^3+6.282\,t^2$ is between 0.0000015 and 0.0000023, so that the constant term should be reduced by 2 in the sixth decimal place. We now have

$$11.154508\,t = 0.006151+, \quad t = 0.0005514+,$$

with doubt only as to whether the last figure of t should be 4 or 5.

Example 1. Find the root between 1 and 2, correct to seven decimal places, of $x^3 + 4\,x^2 - 7 = 0$.

See p. 118. The figure in the fourth decimal place is evidently 2. Thus

$$x = 1.164 + y, \quad 0.0002 < y < 0.0003, \quad y^3 + 7.492\,y^2 + \cdots = 0,$$

$$
\begin{aligned}
0.000000299 &< y^3 + 7.492\,y^2 < 0.000000675,\\
0.003316381 &< 13.376688\,y \ \ < 0.003316757,\\
0.00024792 \ &< \qquad y \qquad < 0.00024795.
\end{aligned}
$$

Hence $x = 1.1642479+$, in which all of the figures are correct. But this work may be abridged. The sum of the terms in y^3 and y^2 has its first significant figure in the seventh decimal place, as shown by $7.5\ (0.0003)^2$. Hence, returning to the final numbers in our transformation scheme above, we carry the division of 0.0033170 by 13.376688 until we reach a remainder whose sign is in doubt in view of the doubt on the seventh decimal place of the dividend. Doubt would here first arise

1	4	0	-7	1
	1	5	5	
1	5	5	-2	
	1	6		
1	6	11		
	1			
1	7			0.1
	0.1	0.71	1.171	
1	7.1	11.71	-0.829	
	0.1	0.72		
1	7.2	12.43		
	0.1			
1	7.3			0.06
	0.06	0.4416	0.772296	
1	7.36	12.8716	-0.056704	
	0.06	0.4452		
1	7.42	13.3168		
	0.06			
1	7.48			0.004
	0.004	0.029936	0.053386944	
1	7.484	13.346736	-0.003317056	
	0.004	0.029952		
1	7.488	13.376688		
	0.004			
1	7.492			

in the case of the figure 9 in the seventh decimal place of the quotient; but this doubt is removed by noting that the correction to be subtracted from the seventh decimal place of the dividend is a figure between 2 and 7 (as shown by the above examination of the terms in y^3 and y^2).

EXAMPLE 2. Find the root between -4 and -3, correct to seven decimal places, of the equation in Ex. 1.

Using the multipliers -4, $+0.6$, $+0.008$, we find that $x = -4 + 0.608 + y$ where

$$y^3 - 6.176\,y^2 + 7.380992\,y - 0.004556288 = 0.$$

Thus y just exceeds 0.0006. The sum of the terms in y^3 and y^2 is -0.000002 to six decimal places. Carrying the division of 0.004558 by 7.381 until the sign of the remainder is in doubt, on account of the doubt in the sixth decimal place, we

get $y = 0.0006175$, with the slight doubt due to the approximate value of the divisor and that of the y^2 term. Since the cube of 6.176 is just less than 235.6 (as shown by logarithms), the sum of the terms in y^3 and y^2 is -0.000002356 to nine decimal places. Carrying out the division of 0.0045586644 by the exact coefficient of y, we get $y = 0.0006176$, correct to seven decimal places. Hence $x = -3.3913823$.

EXERCISES

1. Find to 7 decimals the root of $x^3 + 4x^2 - 7 = 0$ between $-1, -2$.
2. Find to 7 decimals all the roots of $x^3 - 7x - 7 = 0$.

Find to 5 decimals all the real roots of

3. $x^3 + 2x + 20 = 0$.
4. $x^3 + 3x^2 - 2x - 5 = 0$.
5. $x^3 + x^2 - 2x - 1 = 0$.
6. $x^4 + 4x^3 - 17.5x^2 - 18x + 58.5 = 0$.
7. $x^4 - 11727x + 40385 = 0$. (G. H. Darwin.)
8. Find to 8 decimals the root between 2 and 3 of $x^3 - x - 9 = 0$ by making only three transformations.

5.† Without the intermediation of the idea of division by $x - h$, we may show directly that the process of § 4 yields the correct transformed equation. For simplicity, we take a cubic equation

$$f(x) \equiv ax^3 + bx^2 + cx + d = 0.$$

Our process was as follows:

$$\begin{array}{lll|l|l}
a & b & c & d & \underline{h|} \\
 & ah & ah^2 + bh & ah^3 + bh^2 + ch & \\
\hline
a & ah + b & ah^2 + bh + c & ah^3 + bh^2 + ch + d & \\
 & ah & 2ah^2 + bh & = f(h) & \\
\hline
a & 2ah + b & 3ah^2 + 2bh + c = f'(h) & & \\
 & ah & & & \\
\hline
a & 3ah + b & = \frac{1}{2} f''(h) & &
\end{array}$$

Hence the transformed equation is

$$\tfrac{1}{6} f'''(h) p^3 + \tfrac{1}{2} f''(h) p^2 + f'(h) p + f(h) = 0.$$

The terms of the left member, read in reverse order, are those of Taylor's formula for the expansion of $f(h + p)$. Hence the above process yields the equation obtained from $f(x) = 0$ by setting $x = h + p$.

6.† **Numerical Cubic Equations.** After finding a real root $r \neq 0$ of

$$f(x) \equiv x^3 + bx^2 + cx + d = 0,$$

we may obtain the remaining roots r_1 and r_2 from

$$r_1 + r_2 = -b - r, \quad r_1 r_2 = \frac{-d}{r} = r^2 + br + c.$$

We have

(2) $$(r_1 - r_2)^2 \equiv (r_1 + r_2)^2 - 4\,r_1 r_2 = b^2 - 4\,c - 2\,br - 3\,r^2.$$

Thus $r_1 - r_2$ is either real or a pure imaginary. Making use also of $r_1 + r_2$, we shall have the real or imaginary expressions of r_1, r_2. As it would be laborious to compute the right member of (2), we may make use of a device. We have

$$(r_1 - r_2)^2 = b^2 - 3\,c - f'(r).$$

The value of $f'(r)$ for the approximate value of r obtained at any stage of Horner's process is the coefficient preceding the last one in the next transformed equation (§ 5).

Example. Let $f(x) = x^3 + 4\,x^2 - 7$. By Ex. 1, p. 117,

$$f'(1.164) = 13.376688.$$

If we continue Horner's process, using the multiplier $m = 0.000248$, and retaining only six decimal places, we see that we must twice add $7.492\,m = 0.001858$ to the preceding f' to get

$$f'(r) = 13.380404, \quad r = 1.164248.$$

But this continuation of Horner's process is unnecessary. Using $f'''(x) = 6$ and the work on p. 118, we have

$$f'(x + m) = f'(x) + mf''(x) + 3\,m^2, \quad \tfrac{1}{2}f''(1.164) = 7.492,$$
$$f'(r) = 13.37\ \ldots + 2\,m\,(7.492) + 0.0000002 = 13.3804042.$$

Hence we get

$$(r_1 - r_2)^2 = 2.6195958, \qquad r_1 - r_2 = 1.6185165,$$
$$r_1 + r_2 = -5.1642479, \quad r_1 = -1.7728657, \quad r_2 = -3.3913822.$$

7†. Numerical Quartic Equations. Let

$$f(x) \equiv x^4 + bx^3 + cx^2 + dx + e = 0$$

have two distinct real roots r and s. When these are found approximately by Horner's process, we get at the same time $f'(r)$, $f'(s)$, approximately. Call the remaining roots r_1 and r_2. Then

$$r_1 + r_2 = -b - r - s,$$
$$r_1 r_2 = c - (r + s)(r_1 + r_2) - rs = c + b(r + s) + r^2 + rs + s^2,$$
$$(r_1 - r_2)^2 = b^2 - 4\,c - 2\,b(r + s) - 3\,r^2 - 2\,rs - 3\,s^2,$$
$$(r_1 - r_2)^2(b + 2\,r + 2\,s) = -(r_1 - r_2)^2(2\,r_1 + 2\,r_2 + b)$$
$$= b^3 - 4\,bc - 8\,c(r + s) + b(-7\,r^2 - 10\,rs - 7\,s^2) - 6\,r^3 - 10\,r^2s - 10\,rs^2 - 6\,s^3.$$

To the second member add the product of 10 by

$$r^3 + r^2s + rs^2 + s^3 + b(r^2 + rs + s^2) + c(r + s) + d = \frac{f(r) - f(s)}{r - s} = 0.$$

Hence

$$(r_1 - r_2)^2(b + 2\,r + 2\,s) = b^3 - 4\,bc + 8\,d + f'(r) + f'(s).$$

From this equation we get $r_1 - r_2$ and then find r_1 and r_2, approximately.

EXERCISES †

1. After finding one of the real roots of the cubic equations in Exs. 2, 3, 4, 5, 8, p. 119, find the remaining roots by § 6.
2. Treat the quartic equations in Exs. 6, 7, p. 119, by § 7.

Find two and then all of the roots of

3. $x^4 + 12\,x + 7 = 0.$
4. $x^4 - 80\,x^3 + 1998\,x^2 - 14937\,x + 5000 = 0.$

8.† Gräffe's Method. First, let all of the n roots $x_1, \ldots, x_n$ be real and distinct numerically. Choose the notation so that x_1 exceeds x_2 numerically and x_2 exceeds x_3 numerically, etc. In

$$(3) \qquad \Sigma x_1^m = x_1^m\left(1 + \frac{x_2^m}{x_1^m} + \frac{x_3^m}{x_1^m} + \cdots\right),$$

each fraction approaches zero as m increases, so that x_1^m is an approximate value of Σx_1^m if m is sufficiently large. Similarly,

$$(4) \qquad \Sigma x_1^m x_2^m = x_1^m x_2^m\left(1 + \frac{x_3^m}{x_1^m} + \frac{x_3^m}{x_2^m} + \cdots + \frac{x_3^m x_4^m}{x_1^m x_2^m} + \cdots\right),$$

so that $x_1^m x_2^m$ is an approximate value of $\Sigma x_1^m x_2^m$ for m large. Now $x_1^m, \ldots, x_n^m$ are the roots of

$$(5) \qquad y^n - \Sigma x_1^m \cdot y^{n-1} + \Sigma x_1^m x_2^m \cdot y^{n-2} - \cdots = 0.$$

As illustrated in the examples below, it is quite easy to form this equation (5) for values of m which are the successive powers of 2. After obtaining the equation in which m is sufficiently large, we divide each coefficient by the preceding coefficient and obtain approximate values of the negatives of $x_1^m, x_2^m, \ldots$. Indeed, the coefficients are approximately

$$1,\ -x_1^m,\ x_1^m x_2^m,\ -x_1^m x_2^m x_3^m,\ \ldots.$$

Example 1. For $x^3 + x^2 - 2\,x - 1 = 0$, we first form the cubic equation whose roots are the squares of the roots x_1, x_2, x_3 of the given equation. To this

end, we transpose the terms x^2, -1, of even degree, square, replace x^2 by y, and get*

$$y^3 - 5y^2 + 6y - 1 = 0,$$

whose roots are $y_1 = x_1^2$, $y_2 = x_2^2$, $y_3 = x_3^2$. Repeating the operation, we get

$$z^3 - 13z^2 + 26z - 1 = 0, \quad v^3 - 117v^2 + 650v - 1 = 0,$$

with the roots $z_1 = y_1^2, \ldots,$ and $v_1 = z_1^2, \ldots.$ Hence the roots of the v-cubic are the 8th powers of x_1, x_2, x_3. By logarithms, the 8th roots of 117, $\frac{650}{117}$, $\frac{1}{650}$ (the approximate values of x_1^8, x_2^8, x_3^8) are 1.813, 1.239, 0.4450, which are therefore approximate numerical values of x_1, x_2, x_3. The next step gives the equation

$$w^3 - 12389w^2 + 422266w - 1 = 0.$$

The 16th roots of 12389, etc., are -1.80225, 1.24676, -0.44504, to which the proper signs have now been prefixed (their product being positive and sum being -1).

Instead of repeating the process, we may now obtain as follows the values of the roots correct to five decimal places. We had the logarithms of the last approximations to the roots and hence see at once that $(x_3/x_2)^{16}$ affects only the 8th decimal place and that $(x_3/x_1)^{16}$ is still smaller. The coefficient of w is $\Sigma x_1^{16}x_2^{16}$, whose expression (4) involves only the first three terms. Hence

$$x_1^{16}x_2^{16} = 422266,$$

correct to 7 decimal places. The reciprocal is x_3^{16}, whence $x_3 = -0.44504$ to 5 decimal places. By the approximate values of x_1 and x_2 from the w-cubic, $(x_2/x_1)^{16} = 0.002751$. Thus

$$1.002751\, x_1^{16} = 12389 = \Sigma x_1^{16},$$

whence $x_1 = -1.80194$ to 5 decimal places. By the displayed equations,

$$x_2^{16} = \frac{422266 \times 1.002751}{12389}, \quad x_2 = 1.24698.$$

We have now found each root correct to five decimal places. As a check, note that the roots are (Ch. VIII, § 3, § 8)

$$2\cos\frac{2\pi}{7}, \quad 2\cos\frac{4\pi}{7}, \quad 2\cos\frac{6\pi}{7}.$$

The above process requires modification if several of the largest roots are equal or approximately equal numerically. If x_1 and x_2 are approximately equal, but sufficiently different from $x_3, \ldots, x_n$, numerically, an approximate value of x_1^m is $\frac{1}{2}\Sigma x_1^m$.

Next, consider a cubic equation with two conjugate imaginary roots

* We may use symmetric functions: $\Sigma y_1 = \Sigma x_1^2 = (\Sigma x_1)^2 - 2\Sigma x_1x_2 = 5$, etc.

x_2 and x_3, whose modulus (Ch. II, § 8) is r, and a real root x_1 numerically greater than r. Then the real number

$$\frac{x_2{}^m}{x_1{}^m} + \frac{x_3{}^m}{x_1{}^m}$$

is numerically less than or equal to the sum

$$2\left(\frac{\text{mod. } x_2}{\pm x_1}\right)^m$$

of the moduli of its two parts, and hence approaches zero as m increases. Thus, by (3), an approximate value of $x_1{}^m$ is $\Sigma x_1{}^m$.

EXAMPLE 2. For $x^3 - 2x - 2 = 0$, $x_1 > 1.7$, $x_2 x_3 = r^2 = 2/x_1$. Since $2 < (1.7)^3$, $r < 1.7 < x_1$. Forming the equation whose roots are the squares of the roots of the x-cubic, that whose roots are the fourth powers, etc., we get

$$\begin{aligned} y^3 - 4y^2 + 4y - 4 &= 0, \\ z^3 - 8z^2 - 16z - 16 &= 0, \\ v^3 - 96v^2 - 256 &= 0. \end{aligned}$$

Thus x_1 is approximately

$$\sqrt[8]{96} = 1.7692 \ldots .$$

By two more steps, we get

$$x_1 = \sqrt[32]{85032960} = 1.769293,$$

correct to six decimal places.

For a cubic equation in which $x_1 < r$, we employ the equation in X obtained by setting $x = 1/X$. Its root $1/x_1$ exceeds numerically the modulus $1/r$ of the imaginary roots $1/x_2$, $1/x_3$. Hence the equation in X is of the type last discussed.

EXERCISES †

1. The equation whose roots are the 8th powers of the roots x_1, x_2, x_3 of $x^3 - 4x^2 - x + 3 = 0$ is

$$w^3 - 74474w^2 + 46213w - 6561 = 0.$$

Dividing the negative of each coefficient by the preceding coefficient and extracting the 8th root of each quotient, we get 4.06443, 0.94, 0.78. The first is a good approximation to x_1. The last two are approximately equal and hence not good approximations to $-x_2$, x_3. To avoid this inconvenience, add unity to each root (*i.e.*, replace x by $X - 1$). Treat the equation in X and so obtain good approximations to x_1, x_2, x_3.

Treat by the present methods

2. $x^3 - 2x - 5 = 0$. 3. $x^3 - 2x^2 - 2 = 0$. 4. $x^3 + 4x^2 - 7 = 0$.
5. $x^3 + 2x + 20 = 0$.

For further details on the determination of imaginary roots by this method, see Encke, *Crelle's Journal*, vol. 22 (1841), p. 193; and examples by G. Bauer, *Vorlesungen über Algebra*, 1903, p. 244; and C. Runge, *Praxis der Gleichungen*, 1900, p. 157.

9.† To determine the imaginary roots of an equation $f(z) = 0$ with real coefficients, expand $f(x + yi)$ by Taylor's Theorem; we get

$$f(x) + f'(x)\,yi - f''(x)\,\frac{y^2}{1\cdot 2} - f'''(x)\,\frac{y^3 i}{1\cdot 2\cdot 3} + \cdots = 0.$$

Since x and y are to be real, and $y \neq 0$,

$$(6)\qquad \begin{cases} f(x) - f''(x)\,\dfrac{y^2}{1\cdot 2} + f''''(x)\,\dfrac{y^4}{1\cdot 2\cdot 3\cdot 4} - \cdots = 0, \\ f'(x) - f'''(x)\,\dfrac{y^2}{1\cdot 2\cdot 3} + f^{(5)}(x)\,\dfrac{y^4}{5!} - \cdots = 0. \end{cases}$$

By eliminating y^2 between these two equations, we obtain an equation $E(x) = 0$, whose real roots x may be found by one of the preceding methods. In general the next to the final step of the elimination gives y^2 as a rational function of x, so that each real x which yields a positive real value of y^2 furnishes a pair of imaginary roots $x \pm yi$ of $f(z) = 0$. But if there are several pairs of imaginary roots with the same real part x, the equation in y^2 used in the final step of the elimination will be of degree greater than unity in y^2.

Example. For $f(z) = z^4 - z + 1$, equations (6) are

$$x^4 - x + 1 - 6x^2y^2 + y^4 = 0, \quad 4x^3 - 1 - 4xy^2 = 0.$$

Thus

$$y^2 = x^2 - \frac{1}{4x}, \quad -4x^6 + x^2 + \frac{1}{16} = 0.$$

The cubic equation in x^2 has the single real root

$$x^2 = 0.528727, \quad x = \pm 0.72714.$$

Then $y^2 = 0.87254$ or 0.184912, and

$$z = x + yi = 0.72714 \pm 0.43001\,i, \quad -0.72714 \pm 0.93409\,i.$$

EXERCISES†

1. For the quartic equation in Ch. V, § 1, eliminate y^2 between equations $X = 0$, $Y = 0$, corresponding to the present pair (6), and get

$$x(x-2)(16\,x^4 - 64\,x^3 + 136\,x^2 - 144\,x + 65) = 0.$$

Show that the last factor has no real root by setting $2\,x = w + 2$ and obtaining $(w^2 + 1)(w^2 + 9) = 0$. Hence find the four sets of real values x, y and hence the four complex roots $x + yi$.

2. If r and s are any two roots of $f(z) = 0$ and we set

$$x = \frac{r+s}{2}, \quad y = \frac{r-s}{2\,i},$$

we have $r = x + yi$, $s = x - yi$, so that $f(x \pm yi) = 0$. Hence $E(x) = 0$ has as its roots the $\frac{1}{2}\,n(n-1)$ half-sums of the roots of $f(z) = 0$ in pairs. If, however, we eliminate x between equations (6) and set $-4\,y^2 = w$, we obtain an equation in w whose roots are the $\frac{1}{2}\,n(n-1)$ squares of the differences of the roots of $f(z) = 0$.

10†. Lagrange's Method.

The root between 1 and 2 of

$$x^3 + 4\,x^2 - 7 = 0$$

may be expressed as a continued fraction. Set $x = 1 + 1/y$. Then *

$$-2\,y^3 + 11\,y^2 + 7\,y + 1 = 0.$$

Since $-2\,y^3 + 11\,y^2$ must be negative, we have $y > 5$. We find by trial that y lies between 6 and 7. Set $y = 6 + 1/z$.

$$\begin{array}{rrrr|l}
-2 & 11 & 7 & 1 & \underline{6} \\
 & -12 & -6 & 6 & \\
\hline
-2 & -1 & 1 & 7 & \\
 & -12 & -78 & & \\
\hline
-2 & -13 & -77 & & \\
 & -12 & & & \\
\hline
-2 & -25 & & &
\end{array}$$

$$-\frac{2}{z^3} - \frac{25}{z^2} - \frac{77}{z} + 7 = 0, \quad 7\,z^3 - 77\,z^2 - 25\,z - 2 = 0.$$

Since $7\,z^3 - 77\,z^2 > 0$, $z > 11$. The value of z lies between 11 and 12. Now

$$x = 1 + \cfrac{1}{6 + \cfrac{1}{z}} = \frac{7\,z+1}{6\,z+1}.$$

* We may of course first set $x = 1 + d$, find the cubic equation in d by our earlier method, and then replace d by $1/y$.

Using $z = 11$, we find that x is just smaller than 1.1642. But z is in fact just greater than 11.3. Using $z = 11.3$, we find that

$$x = \frac{80.1}{68.8} = 1.1642 +.$$

Hence $x = 1.1642$ to four decimal places.

There is a rapid method of evaluating a continued fraction and a means of finding the limits of the error made in stopping the development at a given place. For an extensive account of the theory and applications of continued fractions, see Serret's *Cours d'Algèbre Supérieure,* ed. 4, I, pp. 7–85, 351–368.

CHAPTER XI

Determinants; Systems of Linear Equations

1. In case there is a pair of numbers x and y for which

$$(1) \qquad \begin{cases} a_1x + b_1y = k_1, \\ a_2x + b_2y = k_2, \end{cases}$$

they may be found as follows. Multiply the members of the first equation by b_2 and those of the second equation by $-b_1$, and add the resulting equations. We get

$$(a_1b_2 - a_2b_1)x = k_1b_2 - k_2b_1.$$

Employing the respective multipliers $-a_2$ and a_1, we get

$$(a_1b_2 - a_2b_1)y = a_1k_2 - a_2k_1.$$

The common multiplier of x and y is

$$(2) \qquad a_1b_2 - a_2b_1,$$

which is called a *determinant of the second order* and denoted by the symbol*

$$(2') \qquad \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.$$

The value of the symbol is obtained by cross-multiplication and substraction. Our earlier results now give

$$(3) \qquad \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} x = \begin{vmatrix} k_1 & b_1 \\ k_2 & b_2 \end{vmatrix}, \qquad \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} y = \begin{vmatrix} a_1 & k_1 \\ a_2 & k_2 \end{vmatrix}.$$

We shall call k_1 and k_2 the known terms of our equations (1). Hence, *if D is the determinant of the coefficients of the unknowns, the product of D by any one of the unknowns equals the determinant obtained from D by substituting the known terms in place of the coefficients of that unknown.*

* The symbol for an expression should show explicitly all of the quantities upon whose values the value of the expression depends. Here these are a_1, b_1, a_2, b_2. The advantage of writing these in the symbol (2′) in the order in which they occur in the equations is that the symbol may be written down without an effort of memory by a mere inspection of the given equations.

EXAMPLE. For $2x - 3y = -4$, $6x - 2y = 2$, we have

$$\begin{vmatrix} 2 & -3 \\ 6 & -2 \end{vmatrix} x = \begin{vmatrix} -4 & -3 \\ 2 & -2 \end{vmatrix}, \qquad 14x = 14, \qquad \mathbf{x = 1},$$

$$14y = \begin{vmatrix} 2 & -4 \\ 6 & 2 \end{vmatrix} = 28, \qquad y = 2.$$

EXERCISES

Solve by determinants the systems of equations

1. $8x - y = 34,\quad x + 8y = 53.$
2. $3x + 4y = 10,\quad 4x + y = 9.$
3. $ax + by = a^2,\quad bx - ay = ab.$

4. Verify that, if the determinant (2) is not zero, the values of x and y determined by division from (3) satisfy equations (1).

2. Consider a system of three linear equations

$$(4)\qquad \begin{aligned} a_1x + b_1y + c_1z &= k_1, \\ a_2x + b_2y + c_2z &= k_2, \\ a_3x + b_3y + c_3z &= k_3. \end{aligned}$$

Multiply the members of the first, second and third equations by *

$$(5)\qquad b_2c_3 - b_3c_2, \qquad b_3c_1 - b_1c_3, \qquad b_1c_2 - b_2c_1,$$

respectively and add the resulting equations. We obtain an equation in which the coefficients of y and z are found to be zero, while the coefficient of x is

$$(6)\qquad a_1b_2c_3 - a_1b_3c_2 + a_2b_3c_1 - a_2b_1c_3 + a_3b_1c_2 - a_3b_2c_1.$$

Such an expression is called a *determinant of the third order* and denoted by the symbol

$$(6')\qquad \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

The nine numbers $a_1, \ldots, c_3$ are called the *elements* of the determinant. In the symbol these elements lie in three (horizontal) *rows*, and also in three (vertical) *columns*. Thus a_2, b_2, c_2 are the elements of the second row, while the three c's are the elements of the third column.

* A simple rule for finding these multipliers is given in § 3.

The equation (free of y and z), obtained above, is

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} x = \begin{vmatrix} k_1 & b_1 & c_1 \\ k_2 & b_2 & c_2 \\ k_3 & b_3 & c_3 \end{vmatrix},$$

since the constant member was the sum of the products of the expressions (5) by k_1, k_2, k_3, and hence may be derived from (6) by replacing the a's by the k's. Thus the theorem of § 1 holds here as regards the value of x.

3. Minors. The determinant of the second order obtained by erasing (or covering up) the row and column crossing at a given element of a determinant of the third order is called the *minor* of that element. For example, in the determinant D given by (6′), the minors of a_1, a_2, a_3 are

$$A_1 = \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix}, \qquad A_2 = \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix}, \qquad A_3 = \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix},$$

respectively. The multipliers (5) are therefore A_1, $-A_2$, A_3. Hence the first results obtained in § 2 may be stated as follows:

$$D = a_1A_1 - a_2A_2 + a_3A_3, \tag{7}$$

$$b_1A_1 - b_2A_2 + b_3A_3 = 0, \qquad c_1A_1 - c_2A_2 + c_3A_3 = 0. \tag{8}$$

The minors of b_1, b_2, b_3 in this determinant D are

$$B_1 = a_2c_3 - a_3c_2, \qquad B_2 = a_1c_3 - a_3c_1, \qquad B_3 = a_1c_2 - a_2c_1.$$

Multiply the members of the equations (4) by $-B_1$, B_2, $-B_3$, respectively, and add. In the resulting equation, the coefficients of x and z are seen to equal zero:

$$-a_1B_1 + a_2B_2 - a_3B_3 = 0, \qquad -c_1B_1 + c_2B_2 - c_3B_3 = 0, \tag{9}$$

while the coefficient of y is seen to equal the expression (6):

$$D = -b_1B_1 + b_2B_2 - b_3B_3. \tag{10}$$

Hence the theorem of § 1 holds here for the variable y.

The reader should also verify that, if he uses the multipliers C_1, $-C_2$, C_3, where C_i is the minor of c_i in D, he obtains an equation in which the coefficients of x and y are zero:

$$a_1C_1 - a_2C_2 + a_3C_3 = 0, \qquad b_1C_1 - b_2C_2 + b_3C_3 = 0, \tag{11}$$

while the coefficient of z equals the expression (6):

$$D = c_1C_1 - c_2C_2 + c_3C_3, \tag{12}$$

and then conclude that the theorem of § 1 is true as regards z.

4. Expansion According to the Elements of a Column. Relations (7), (10), (12) are expressed in words by saying that *a determinant of the third order may be expanded according to the elements of any column.* To obtain the expansion, we multiply each element of the column by the minor of the element, prefix the proper sign to the products, and add the signed products. The signs are alternately + and −, as in the diagram.

$$\begin{matrix} + & - & + \\ - & + & - \\ + & - & + \end{matrix}$$

5. Two Columns Alike. *A determinant* is zero if any two of its columns are alike.*

This is evident for a determinant of the second order:

$$\begin{vmatrix} c & c \\ d & d \end{vmatrix} = cd - cd = 0.$$

To prove it for a determinant of the third order, we have only to expand it according to the elements of the column not one of the like columns and to note that each minor is zero, being a determinant of the second order with two columns alike.

EXERCISES

Solve by determinants the systems of equations (expanding a determinant having two zeros in a column according to the elements of that column):

1. $x + y + z = 11,$ $2x - 6y - z = 0,$ $3x + 4y + 2z = 0.$

2. $x + y + z = 0,$ $x + 2y + 3z = -1,$ $x + 3y + 6z = 0.$

3. Noting that A_1, A_2, A_3 of § 3 do not involve a_1, a_2, a_3, we may obtain the first expression (8) from (7) by replacing each a_i by b_i, and the second expression (8) from (7) by replacing each a_i by c_i. Hence (8) are the expansions of

$$\begin{vmatrix} b_1 & b_1 & c_1 \\ b_2 & b_2 & c_2 \\ b_3 & b_3 & c_3 \end{vmatrix} = 0, \qquad \begin{vmatrix} c_1 & b_1 & c_1 \\ c_2 & b_2 & c_2 \\ c_3 & b_3 & c_3 \end{vmatrix} = 0$$

according to the elements of the first column.

4. Prove similarly that (9) and (11) follow from § 5.

* Here and in §§ 6–11 we understand by a determinant one of the second or third order. After determinants of higher orders have been defined, it will be shown that these theorems are true of determinants of any order.

6. Theorem. *A determinant having $a_1 + q_1$, $a_2 + q_2$, . . . as the elements of a column equals the sum of the determinant having a_1, a_2, . . . as the elements of the corresponding column and the determinant having q_1, q_2, . . . as the elements of that column, while the elements of the remaining columns of each determinant are the same as in the given determinant.*

For determinants of the second order, there are only two cases:

$$\begin{vmatrix} a_1 + q_1 & b_1 \\ a_2 + q_2 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} + \begin{vmatrix} q_1 & b_1 \\ q_2 & b_2 \end{vmatrix},$$

$$\begin{vmatrix} b_1 & a_1 + q_1 \\ b_2 & a_2 + q_2 \end{vmatrix} = \begin{vmatrix} b_1 & a_1 \\ b_2 & a_2 \end{vmatrix} + \begin{vmatrix} b_1 & q_1 \\ b_2 & q_2 \end{vmatrix}.$$

For determinants of the third order, one of the three cases is

$$\begin{vmatrix} a_1 + q_1 & b_1 & c_1 \\ a_2 + q_2 & b_2 & c_2 \\ a_3 + q_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} q_1 & b_1 & c_1 \\ q_2 & b_2 & c_2 \\ q_3 & b_3 & c_3 \end{vmatrix}.$$

To prove the theorem we have only to expand the three determinants according to the elements of the column in question (the first column in the first and third illustrations, the second column in the second illustration) and note that the minors are the same for all three determinants. Hence $a_1 + q_1$ is multiplied by the same minor that a_1 and q_1 are multiplied by separately, and similarly for $a_2 + q_2$, etc.

7. Removal of Factors. *A common factor of all of the elements of the same column of a determinant may be divided out of the elements and placed as a factor before the new determinant.*

In other words, if all of the elements of a column are divided by n, the value of the determinant is divided by n. For example,

$$\begin{vmatrix} na_1 & b_1 \\ na_2 & b_2 \end{vmatrix} = n\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \qquad \begin{vmatrix} a_1 & nb_1 & c_1 \\ a_2 & nb_2 & c_2 \\ a_3 & nb_3 & c_3 \end{vmatrix} = n\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

Proof is made by expanding the determinants according to the elements of the column in question.

8. Theorem. *A determinant is not changed in value if we add to the elements of any column the products of the corresponding elements of another column by the same number.*

For example,
$$\begin{vmatrix} a_1 + nb_1 & b_1 \\ a_2 + nb_2 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix},$$
as follows from the first relation in § 6. Similarly, by the third,
$$\begin{vmatrix} a_1 + nb_1 & b_1 & c_1 \\ a_2 + nb_2 & b_2 & c_2 \\ a_3 + nb_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + n \begin{vmatrix} b_1 & b_1 & c_1 \\ b_2 & b_2 & c_2 \\ b_3 & b_3 & c_3 \end{vmatrix},$$
in which the last determinant is zero by § 5.

In general, let $a_1, a_2, \ldots$ be the elements to which we add the products of the elements $b_1, b_2, \ldots$ by n. We apply § 6 with $q_1 = nb_1$, $q_2 = nb_2, \ldots$. Thus the modified determinant equals the sum of the initial determinant and a determinant having $b_1, b_2, \ldots$ in one column and $nb_1, nb_2, \ldots$ in another column. But the latter determinant equals (§ 7) the product of n by a determinant with two columns alike and hence is zero (§ 5).

Example. Multiplying the elements of the last column by 2 and adding the products to the elements of the second column, we get
$$\begin{vmatrix} 1 & -2 & 1 \\ 1 & 2 & 3 \\ 6 & 4 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 1 \\ 1 & 8 & 3 \\ 6 & 10 & 3 \end{vmatrix} = \begin{vmatrix} 0 & 0 & 1 \\ -2 & 8 & 3 \\ 3 & 10 & 3 \end{vmatrix} = \begin{vmatrix} -2 & 8 \\ 3 & 10 \end{vmatrix} = -44.$$
For the next step, we have multiplied the elements of the third column by -1 and added the products to the elements of the first column. Expanding the third determinant according to the elements of the third column, we note that two of the minors are zero (having a row of zeros), and hence obtain the determinant of the second order written above. The last step is simplified by use of § 10.

9. Interchange of Rows and Columns. *A determinant is not altered if in its symbol we take as the elements of the first, second, . . . rows the elements (in the same order) which formerly appeared in the first, second, . . . columns:*
$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix},$$
$$D \equiv \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \equiv \Delta.$$

The proof is evident by inspection for the case of determinants of the second order. For those of the third order, we expand Δ and find that its six terms are those in the expansion (6) of D.

10. Expansion According to the Elements of a Row. To prove that determinant D, given by (6′), *may be expanded according to the elements of any row* (say the second *):

$$D = -a_2A_2 + b_2B_2 - c_2C_2,$$

with the same rule of signs as in § 4, we note that (§ 9)

$$D = \Delta = -a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + b_2 \begin{vmatrix} a_1 & a_3 \\ c_1 & c_3 \end{vmatrix} - c_2 \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix},$$

since Δ can be expanded according to the elements of its second column. After interchanging the rows and columns in these three determinants of the second order, we have the minors A_2, B_2, C_2 of a_2, b_2, c_2 in D.

Example. The third determinant in the Example of § 8 is best evaluated by expanding it according to the elements of its first row, since two of its elements are zero. Indeed, we obtain $+1$ multiplied by its minor.

11. Theorem. *A determinant is not changed in value if we add to the elements of any row the products of the corresponding elements of another row by the same number.*

We shall show that D, given by (6′), equals

$$D' \equiv \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 + na_1 & b_2 + nb_1 & c_2 + nc_1 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

Now $D = \Delta$, where Δ is given in § 9. By § 8,

$$\Delta = \begin{vmatrix} a_1 & a_2 + na_1 & a_3 \\ b_1 & b_2 + nb_1 & b_3 \\ c_1 & c_2 + nc_1 & c_3 \end{vmatrix}.$$

Interchanging the rows and columns of Δ, we get D'. Hence

$$D' = \Delta = D.$$

* While for concreteness we have here (and in § 11) treated but one of several cases, the proof is such that it applies to all the cases.

EXERCISES

1. Evaluate the numerical determinant in § 8 by removing the factor 2 from the second column and then getting a determinant with two zeros in the second row.

Solve the systems of equations (by removing, if possible, integral factors from a column and reducing each determinant to one with two zeros in a row before expanding it):

2. $x - 2y + z = 12,$
$x + 2y + 3z = 48,$
$6x + 4y + 3z = 84.$

3. $3x - 2y = 7,$
$3y - 2z = 6,$
$3z - 2x = -1.$

Factor a single determinant, and solve

4. $x + y + z = 1,$
$ax + by + cz = k,$
$a^2x + b^2y + c^2z = k^2.$

5. $ax + by + cz = k,$
$a^2x + b^2y + c^2z = k^2,$
$a^4x + b^4y + c^4z = k^4.$

6. Obtain in its simplest form the value of x from

$$ax + y + z = a - 3,$$
$$x + ay + z = -2,$$
$$x + y + az = -2.$$

7. Deduce the case $n = 2$ of § 7 at once from § 6, by taking $q_i = a_i$.
8. Give the proof in § 10 when the third row is used.
9. Give the proof in § 11 for a new case.
10. A determinant of the third order is zero if two rows are alike.
11. Hence prove that $D' = D$ in § 11 by expanding D' according to the elements of its second row.
12. Prove the theorem about rows corresponding to that in § 6.
13. From Ex. 12 deduce Ex. 11.

12. Definition of a Determinant of Order n. In the six terms of the expression (6), which was defined to be the general determinant of order 3, the letters a, b, c were always written in this sequence, while the subscripts are the six possible arrangements of the numbers 1, 2, 3. The first term $a_1b_2c_3$ shall be called the *diagonal term*,* since it is the product of the elements in the main diagonal running from the upper left hand corner to the lower right hand corner of the symbol for the determinant. The subscripts in the term $-a_1b_3c_2$ are derived from those of the diagonal term by interchanging 2 and 3, and the minus sign is to be associated with the fact that an odd number (here one) of interchanges of subscripts were used. To obtain the arrangement 2, 3, 1 of the subscripts in the

* Sometimes called the leading term.

term $+a_2b_3c_1$ from the natural order 1, 2, 3 (in the diagonal term), we may first interchange 1 and 2, obtaining 2, 1, 3 and then interchange 1 and 3; an even number (two) of interchanges of subscripts were used and the sign of the term is plus.

EXERCISES

1. Show that a like result holds for the last three terms of (6).
2. Discuss similarly the two terms of a determinant of order 2.

While the arrangement 1, 3, 2 was obtained from 1, 2, 3 by one interchange (2, 3), we may obtain it by applying in succession the three interchanges (1, 2), (1, 3), (1, 2), and in many new ways. To show that the number of interchanges which will produce the final arrangement 1, 3, 2 is odd in every case, note that any interchange (the possible ones being the three just listed) changes the sign of the product

$$P = (x_1 - x_2)(x_1 - x_3)(x_2 - x_3),$$

where the x's are arbitrary variables. Thus a succession of k interchanges yields P or $-P$ according as k is even or odd. Starting with the arrangement 1, 2, 3 and applying k successive interchanges, suppose that we obtain the final arrangement 1, 3, 2. But if in P we replace the subscripts 1, 2, 3 by 1, 3, 2, respectively, *i.e.*, if we interchange 2 and 3, we obtain $-P$. Hence k is odd.

Consider the corresponding question for n variables. Form the product of all of the differences $x_i - x_j$ $(i < j)$ of the variables:

$$\begin{aligned} P = (x_1 - x_2)(x_1 - x_3) &\ldots (x_1 - x_n) \\ \cdot (x_2 - x_3) &\ldots (x_2 - x_n) \\ \ldots\ldots &\ldots\ldots \\ &\cdot (x_{n-1} - x_n). \end{aligned}$$

Interchange any two subscripts i and j. The factors which involve neither i nor j are unaltered. The factor $\pm(x_i - x_j)$ involving both is changed in sign. The remaining factors may be paired to form the products

$$\pm(x_i - x_k)(x_j - x_k) \qquad (k = 1, \ldots, n;\ k \neq i,\ k \neq j).$$

Such a product is unaltered. Hence P is changed in sign.

Suppose that an arrangement $i_1, i_2, \ldots, i_n$ can be obtained from $1, 2, \ldots, n$ by using a successive interchanges and also by b successive interchanges. Make these interchanges on the subscripts in P; the

resulting functions equal $(-1)^a P$ and $(-1)^b P$, respectively. But the resulting functions are identical since either can be obtained at one step from P by replacing the subscript 1 by i_1, 2 by i_2, . . . , n by i_n. Hence

$$(-1)^a P \equiv (-1)^b P,$$

so that a and b are both even or both odd.

We define a determinant of order 4 to be

$$\text{(13)} \qquad \begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix} = \sum_{(24)} \pm a_q b_r c_s d_t,$$

where q, r, s, t is any one of the 24 arrangements of 1, 2, 3, 4, and the sign of the corresponding term is + or − according as an even or odd number of interchanges are needed to derive this arrangement q, r, s, t from 1, 2, 3, 4. Although different numbers of interchanges will produce the same arrangement q, r, s, t from 1, 2, 3, 4, these numbers are all even or all odd, as just proved, so that the sign is fully determined.

We have seen that the analogous definitions of determinants of orders 2 and 3 lead to our earlier expressions (2) and (6).

We will have no difficulty in extending the definition to a determinant of general order n as soon as we decide upon a proper notation for the n^2 elements. The subscripts 1, 2, . . . , n may be used as before to specify the rows. But the alphabet does not contain n letters with which to specify the columns. The use of e', e'', . . . , $e^{(n)}$ for this purpose would conflict with the notation for derivatives and besides be very awkward when exponents are used. It is customary in mathematical journals and scientific books (a custom not always followed in introductory text books, to the distinct disadvantage of the reader) to denote the n letters used to distinguish the n columns by $e_1, e_2, \ldots, e_n$ (or some other letter with the same subscripts) and to prefix (but see § 13) such a subscript by the subscript indicating the row. The symbol for the determinant is therefore

$$\text{(14)} \qquad D = \begin{vmatrix} e_{11} & e_{12} & \cdots & e_{1n} \\ e_{21} & e_{22} & \cdots & e_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ e_{n1} & e_{n2} & \cdots & e_{nn} \end{vmatrix}.$$

By definition * this shall mean the sum of the $n!$ terms

$$(14') \qquad (-1)^i e_{i_1 1} e_{i_2 2} \dots e_{i_n n}$$

in which $i_1, i_2, \dots, i_n$ is an arrangement of $1, 2, \dots, n$, derived from $1, 2, \dots, n$ by i interchanges. For example, if we take $n = 4$ and write a_j, b_j, c_j, d_j for $e_{j1}, e_{j2}, e_{j3}, e_{j4}$, the symbol (14) becomes (13) and the general term (14′) becomes $(-1)^i a_{i_1} b_{i_2} c_{i_3} d_{i_4}$, the general term of the second member of (13).

EXERCISES

1. Give the six terms involving a_2 in the determinant (13).
2. What are the signs of $a_3b_5c_2d_1e_4$, $a_5b_4c_3d_2e_1$ in a determinant of order five?
3. The arrangement 4, 1, 3, 2 may be obtained from 1, 2, 3, 4 by use of the two successive interchanges (1, 4), (1, 2), and also by use of the four successive interchanges (1, 4), (1, 3), (1, 2), (2, 3).
4. Write out the six terms of (14) for $n = 3$, rearrange the factors of each term so that the new first subscripts shall be in the order 1, 2, 3, and verify that the resulting six terms are those of the expansion of D' in § 13 for $n = 3$.

13. Interchange of Rows and Columns. *Determinant* (14) *equals*

$$D' = \begin{vmatrix} e_{11} & e_{21} & \dots & e_{n1} \\ e_{12} & e_{22} & \dots & e_{n2} \\ \cdot & \cdot & \cdot & \cdot \\ e_{1n} & e_{2n} & \dots & e_{nn} \end{vmatrix}.$$

Without altering (14′), we may rearrange its factors so that the first subscripts shall appear in the order $1, 2, \dots, n$, and get

$$(-1)^i e_{1k_1} e_{2k_2} \dots e_{nk_n}.$$

Since this can be done by i interchanges of the letters e (corresponding to the i interchanges by which the first subscripts $i_1, \dots, i_n$ were derived from $1, \dots, n$), the new second subscripts $k_1, \dots, k_n$ are derived from the old second subscripts $1, \dots, n$ by i interchanges. The resulting signed product is therefore a term of D'. Hence $D = D'$.

* We may define a determinant of order n by mathematical induction from $n - 1$ to n, using the first equation in § 17. The next step would be to prove that the present definition holds as a theorem.

14. Interchange of Two Columns. *A determinant is changed in sign by the interchange of any two of its columns.*

Let Δ be the determinant derived from (14) by the interchange of the rth and sth columns. The expansion of Δ is therefore obtained from that of D by interchanging r and s in the series of second subscripts of each term (14′) of D. Interchange the rth and sth letters e to restore the second subscripts to their natural order. Since the first subscripts have undergone an interchange, the negative of any term of Δ is a term of D, and $\Delta = -D$.

15. Interchange of Two Rows. *A determinant D is changed in sign by the interchange of any two rows.*

Let Δ be the determinant obtained from D by interchanging the rth and sth rows. By interchanging the rows and columns in D and in Δ, we get two determinants D' and Δ', either of which may be derived from the other by the interchange of the rth and sth columns. Hence, by §§ 13, 14,

$$\Delta = \Delta' = -D' = -D.$$

16. Two Rows or Two Columns Alike. *A determinant is zero if any two of its rows or any two of its columns are alike.*

For, by the interchange of the two like rows or two like columns, the determinant is evidently unaltered, and yet must change in sign by §§ 14 15. Hence $D = -D$, $D = 0$.

17. Expansion. *A determinant can be expanded according to the elements of any row or any column.*

Let E_{ij} be the minor of e_{ij} in D, given by (14). Thus E_{ij} is the determinant of order $n - 1$ obtained by erasing the ith row and the jth column (crossing at e_{ij}). We first prove that

$$D = e_{11}E_{11} - e_{21}E_{21} + e_{31}E_{31} - \cdots + (-1)^{n-1}e_{n1}E_{n1},$$

so that D can be expanded according to the elements of its first column. The terms of D with the factor e_{11} are of the form

$$(-1)^i e_{11}e_{i_2 2} \ldots e_{i_n n},$$

where $i_2, \ldots, i_n$ is an arrangement of $2, \ldots, n$ derived from the latter by i interchanges. Removing from each term the factor e_{11}, and adding the quotients, we obtain the $(n - 1)!$ properly signed terms of E_{11}.

Let Δ be the determinant obtained from D by interchanging the first and second rows. As just proved, the total coefficient of e_{21} in Δ is the minor

$$\begin{vmatrix} e_{12} & e_{13} & \cdots & e_{1n} \\ e_{32} & e_{33} & \cdots & e_{3n} \\ \cdot & \cdot & \cdot & \cdot \\ e_{n2} & e_{n3} & \cdots & e_{nn} \end{vmatrix}$$

of e_{21} in Δ. Now this minor is identical with E_{21}. But $\Delta = -D$ (§ 15). Hence the total coefficient of e_{21} in D equals $-E_{21}$.

Similarly, the coefficient of e_{31} is E_{31}, etc.

To obtain the expansion of D according to the elements of its kth column, where $k > 1$, we consider the determinant δ derived from D by moving the kth column over the earlier columns until it becomes the new first column.

Since this may be done by $k - 1$ interchanges of adjacent columns, $\delta = (-1)^{k-1}D$. The minors of the elements $e_{1k}, \ldots, e_{nk}$ in the first column of δ are evidently the minors $E_{1k}, \ldots, E_{nk}$ of $e_{1k}, \ldots, e_{nk}$ in D. Hence, by the earlier result,

$$D = \sum_{j=1}^{n} (-1)^{j+k} e_{jk} E_{jk} \qquad (k = 1, \ldots, n). \tag{15}$$

Applying this result to the equal determinant D' of § 13, and changing the summation index from j to k, we get

$$D = \sum_{k=1}^{n} (-1)^{j+k} e_{jk} E_{jk} \qquad (j = 1, \ldots, n). \tag{16}$$

This gives the expansion of D according to the elements of the jth row. One decided advantage of the double subscript notation is the resulting simplicity of the last two expansions. Of course the sign may also be found by counting spaces as in § 4.

18. The theorems in §§ 6–8, 11 now follow for determinants of order n. Indeed, the proofs were so worded that they now apply, since the auxiliary theorems used have been extended (§§ 13, 16, 17) to determinants of order n.

EXERCISES

1. Prove the theorem of § 15 by the direct method of § 14.

2.
$$\begin{vmatrix} b+c & c+a & a+b \\ b_1+c_1 & c_1+a_1 & a_1+b_1 \\ b_2+c_2 & c_2+a_2 & a_2+b_2 \end{vmatrix} = 2 \begin{vmatrix} a & b & c \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix}.$$

By reducing to a determinant of order 3, etc., prove that

3.
$$\begin{vmatrix} 2 & -1 & 3 & -2 \\ 1 & 7 & 1 & -1 \\ 3 & 5 & -5 & 3 \\ 4 & -3 & 2 & -1 \end{vmatrix} = -42.$$

4.
$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{vmatrix} = 1.$$

5.
$$\begin{vmatrix} a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^3 & b^3 & c^3 & d^3 \\ a^4 & b^4 & c^4 & d^4 \end{vmatrix} = abcd(a-b)(a-c)(a-d)(b-c)(b-d)(c-d).$$

6.
$$\begin{vmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \cdot \begin{vmatrix} e & f \\ g & h \end{vmatrix}. \qquad \text{[use § 6].}$$

7.
$$\begin{vmatrix} a_1e_1+b_1e_2+c_1e_3 & a_1f_1+b_1f_2+c_1f_3 & a_1g_1+b_1g_2+c_1g_3 \\ a_2e_1+b_2e_2+c_2e_3 & a_2f_1+b_2f_2+c_2f_3 & a_2g_1+b_2g_2+c_2g_3 \\ a_3e_1+b_3e_2+c_3e_3 & a_3f_1+b_3f_2+c_3f_3 & a_3g_1+b_3g_2+c_3g_3 \end{vmatrix}$$
$$= \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \cdot \begin{vmatrix} e_1 & f_1 & g_1 \\ e_2 & f_2 & g_2 \\ e_3 & f_3 & g_3 \end{vmatrix}.$$

Write out only the 6 of the 27 determinants (§ 6) which are not necessarily zero.

8. Hence verify that the product of two determinants of the same order (2 or 3) is a determinant of like order in which the element of the rth row and cth column is the sum of the products of the elements of the rth row of the first determinant by the corresponding elements of the cth column of the second.

9. Express $(a^2+b^2+c^2+d^2)(e^2+f^2+g^2+h^2)$ as a sum of 4 squares by writing

$$\begin{vmatrix} a+bi & c+di \\ -c+di & a-bi \end{vmatrix} \cdot \begin{vmatrix} e+fi & g+hi \\ -g+hi & e-fi \end{vmatrix}$$

as a determinant of order 2 similar to each factor.

10. If $s_i = \alpha^i + \beta^i + \gamma^i$,

$$\begin{vmatrix} 1 & 1 & 1 \\ \alpha & \beta & \gamma \\ \alpha^2 & \beta^2 & \gamma^2 \end{vmatrix} \cdot \begin{vmatrix} 1 & \alpha & \alpha^2 \\ 1 & \beta & \beta^2 \\ 1 & \gamma & \gamma^2 \end{vmatrix} = \begin{vmatrix} 3 & s_1 & s_2 \\ s_1 & s_2 & s_3 \\ s_2 & s_3 & s_4 \end{vmatrix}.$$

11. Using the Factor Theorem and the diagonal term, prove Ex. 5 and

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \cdot & \cdot & \cdot & \cdot \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix} = \prod_{\substack{i,j=1 \\ i>j}}^{n} (x_i - x_j) = (-1)^{\frac{n(n-1)}{2}} P,$$

where P is given in § 12.

12. With the notations of § 3, and using (7)–(12), prove that

$$\begin{vmatrix} A_1 & -A_2 & A_3 \\ -B_1 & B_2 & -B_3 \\ C_1 & -C_2 & C_3 \end{vmatrix} \cdot \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} D & 0 & 0 \\ 0 & D & 0 \\ 0 & 0 & D \end{vmatrix}.$$

Hence the first determinant equals D^2.

19. Complementary Minors. The determinant D of order 4 in (13) is said to have the *two-rowed complementary minors*

$$M = \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix}, \qquad M' = \begin{vmatrix} c_2 & d_2 \\ c_4 & d_4 \end{vmatrix},$$

since either is obtained by erasing from D all the rows and columns having an element occurring in the other. Similarly, any r-rowed minor of a determinant of order n has a definite complementary $(n-r)$-rowed minor. In particular, any element is regarded as a one-rowed minor and is complementary to its minor.

20. Laplace's Development. *Any determinant D equals the sum of all the products $\pm MM'$, where M is an r-rowed minor having its elements in the first r columns of D, and M' is the minor complementary to M, while the sign is $+$ or $-$ according as an even or odd number of interchanges of rows of D will bring M into the position occupied by the minor M_1 whose elements lie in the first r rows and first r columns of D.*

For $r = 1$, this development becomes the known expansion of D according to the elements of the first column (§ 17); here $M_1 = e_{11}$.

If $r = 2$ and D is the determinant (13) of order 4,

$$D = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \cdot \begin{vmatrix} c_3 & d_3 \\ c_4 & d_4 \end{vmatrix} - \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} \cdot \begin{vmatrix} c_2 & d_2 \\ c_4 & d_4 \end{vmatrix} + \begin{vmatrix} a_1 & b_1 \\ a_4 & b_4 \end{vmatrix} \cdot \begin{vmatrix} c_2 & d_2 \\ c_3 & d_3 \end{vmatrix}$$

$$+ \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} \cdot \begin{vmatrix} c_1 & d_1 \\ c_4 & d_4 \end{vmatrix} - \begin{vmatrix} a_2 & b_2 \\ a_4 & b_4 \end{vmatrix} \cdot \begin{vmatrix} c_1 & d_1 \\ c_3 & d_3 \end{vmatrix} + \begin{vmatrix} a_3 & b_3 \\ a_4 & b_4 \end{vmatrix} \cdot \begin{vmatrix} c_1 & d_1 \\ c_2 & d_2 \end{vmatrix}.$$

The first term of the development is M_1M_1'; the second term is $-MM'$ (in the notations of § 19), and the sign is minus since the interchange of the second and third rows of D brings this M into the position of M_1. The sign of the third term of the development is plus since two interchanges of rows of D bring the first factor into the position of M_1.

If D is the determinant (14), then

$$M_1 = \begin{vmatrix} e_{11} & \cdot & \cdot & \cdot & e_{1r} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ e_{r1} & \cdot & \cdot & \cdot & e_{rr} \end{vmatrix}, \quad M_1' = \begin{vmatrix} e_{r+1\,r+1} & \cdot & \cdot & \cdot & e_{r+1\,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ e_{n\,r+1} & \cdot & \cdot & \cdot & e_{n\,n} \end{vmatrix}.$$

Any term of the product M_1M_1' is of the type

$$(-1)^i e_{i_1 1}\, e_{i_2 2} \ . \ . \ . \ e_{i_r r} \cdot (-1)^j e_{i_{r+1} r+1} \ . \ . \ . \ e_{i_n n},$$

where $i_1, \ldots, i_r$ is an arrangement of $1, \ldots, r$ derived from $1, \ldots, r$ by i interchanges, while $i_{r+1}, \ldots, i_n$ is an arrangement of $r + 1, \ldots, n$ derived by j interchanges. Hence $i_1, \ldots, i_n$ is an arrangement of $1, \ldots, n$ derived by $i + j$ interchanges, so that the above product is a term of D with the proper sign.

It now follows from § 15 that any term of any of the products $\pm MM'$ of the theorem is a term of D. Clearly we do not obtain in this manner the same term of D twice.

Conversely, any term t of D occurs in one of the products $\pm MM'$. Indeed, t contains as factors r elements from the first r columns of D, no two being in the same row, and the product of these is, except perhaps as to sign, a term of some minor M. Thus t is a term of MM' or of $-MM'$. In view of the earlier discussion, the sign of t is that of the corresponding term in $\pm MM'$, where the latter sign is given by the theorem.

21. There is a Laplace development of D in which the r-rowed minors M have their elements in the first r rows of D, instead of in the first r columns as in § 20. To prove this, we have only to apply § 20 to the equal determinant obtained by interchanging the rows and columns of D.

There are more general (but less used) Laplace developments in which the r-rowed minors M have their elements in any chosen r columns (or rows) of D. It is simpler to apply the earlier developments to the determinant $\pm D$ having the elements of the chosen r columns (or rows) in the new first r columns (or rows).

EXERCISES

1. $$\begin{vmatrix} a & b & c & d \\ e & f & g & h \\ 0 & 0 & j & k \\ 0 & 0 & l & m \end{vmatrix} = \begin{vmatrix} a & b \\ e & f \end{vmatrix} \cdot \begin{vmatrix} j & k \\ l & m \end{vmatrix}.$$

2. $$\frac{1}{2}\begin{vmatrix} a & b & c & d \\ e & f & g & h \\ a & b & c & d \\ e & f & g & h \end{vmatrix} = \begin{vmatrix} a & b \\ e & f \end{vmatrix} \cdot \begin{vmatrix} c & d \\ g & h \end{vmatrix} - \begin{vmatrix} a & c \\ e & g \end{vmatrix} \cdot \begin{vmatrix} b & d \\ f & h \end{vmatrix} + \begin{vmatrix} a & d \\ e & h \end{vmatrix} \cdot \begin{vmatrix} b & c \\ f & g \end{vmatrix} = 0.$$

3. Check § 20 by showing that the total number of products of n elements is $C_r{}^n \cdot r!(n-r)! = n!$, where $C_r{}^n$ is the number of combinations of n things r at a time.

For Laplace's development of many special determinants, see Ch. XII.

22. Product of Determinants. The important rule (Ex. 8, p. 140), for expressing the product of two determinants of order n as a determinant of order n is found and proved easily by means of Laplace's development. For brevity we shall take $n = 3$, but the method is seen to apply for any n. We have

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \cdot \begin{vmatrix} e_1 & f_1 & g_1 \\ e_2 & f_2 & g_2 \\ e_3 & f_3 & g_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 & 0 & 0 & 0 \\ a_2 & b_2 & c_2 & 0 & 0 & 0 \\ a_3 & b_3 & c_3 & 0 & 0 & 0 \\ -1 & 0 & 0 & e_1 & f_1 & g_1 \\ 0 & -1 & 0 & e_2 & f_2 & g_2 \\ 0 & 0 & -1 & e_3 & f_3 & g_3 \end{vmatrix}$$

In the determinant of order 6, multiply the elements of the first column by e_1, f_1, g_1 in turn and add the products to the corresponding elements of the fourth, fifth and sixth columns, respectively (and hence introduce zeros in place of the present elements e_1, f_1, g_1). Then multiply the elements of the second column by e_2, f_2, g_2 in turn and add the products to the corresponding elements of the fourth, fifth and sixth columns, respectively. Finally, multiply the elements of the third column by e_3, f_3, g_3 in turn and add as before. The new determinant is

$$\begin{vmatrix} a_1 & b_1 & c_1 & a_1e_1+b_1e_2+c_1e_3 & a_1f_1+b_1f_2+c_1f_3 & a_1g_1+b_1g_2+c_1g_3 \\ a_2 & b_2 & c_2 & a_2e_1+b_2e_2+c_2e_3 & a_2f_1+b_2f_2+c_2f_3 & a_2g_1+b_2g_2+c_2g_3 \\ a_3 & b_3 & c_3 & a_3e_1+b_3e_2+c_3e_3 & a_3f_1+b_3f_2+c_3f_3 & a_3g_1+b_3g_2+c_3g_3 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{vmatrix}.$$

By Laplace's development (or by expansion according to the elements of the last row, etc.), this equals the 3-rowed minor whose elements are the long sums, and written in Ex. 7, p. 140.

23. Systems of Linear Equations. In the n equations

$$(17)\qquad \begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = k_1, \\ \cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = k_n, \end{array}$$

let D denote the determinant of the coefficients of the n unknowns:

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

Let A_{ij} be the minor of a_{ij} in D. Multiply the members of the first equation by A_{11}, those of the second equation by $-A_{21}, \ldots$, those of the nth equation by $(-1)^{n-1}A_{n1}$, and add. The coefficient of x_1 is the expansion of D according to the elements of its first column. The coefficient of x_2 is the expansion, according to the elements of the first column, of a determinant derived from D by replacing a_{11} by $a_{12}, \ldots, a_{n1}$ by a_{n2}, so that this determinant has the first two columns alike and hence is zero. In this manner, we find that

$$(18)\qquad Dx_1 = K_1, \qquad Dx_2 = K_2, \ldots, \qquad Dx_n = K_n,$$

in which (see (β) of § 24) K_i is derived from D by substituting $k_1, \ldots, k_n$ for the elements $a_{1i}, \ldots, a_{ni}$ of the ith column of D. Another proof of (18) follows from

$$Dx_1 = \begin{vmatrix} a_{11}x_1 & a_{12} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1}x_1 & a_{n2} & \ldots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11}x_1 + \cdots + a_{1n}x_n & \ldots & a_{1n} \\ \cdot & \cdot & \cdot \\ a_{n1}x_1 + \cdots + a_{nn}x_n & \ldots & a_{nn} \end{vmatrix} = K_1.$$

We have now extended to any n results proved for $n = 2$ and $n = 3$ in §§ 1–3.

If $D \neq 0$, the unique values of $x_1, \ldots, x_n$ determined by division from (18) actually satisfy equations (17). For instance, the first equation is satisfied since

$$k_1D - a_{11}K_1 - a_{12}K_2 - \cdots - a_{1n}K_n = \begin{vmatrix} k_1 & a_{11} & a_{12} & \ldots & a_{1n} \\ k_1 & a_{11} & a_{12} & \ldots & a_{1n} \\ k_2 & a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ k_n & a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix},$$

as shown by expansion according to the elements of the first row; and the determinant is zero, having two rows alike.

24. Rank of a Determinant. If a determinant D of order n is not zero, it is said to be of rank n. In general, if some r-rowed minor of D is not zero, while every $(r + 1)$-rowed minor is zero, D is said to be of rank r.

For example, a determinant D of order 3 is of rank 3 if $D \neq 0$; of rank 2 if $D = 0$, but some two-rowed minor is not zero; it is of rank 1 if every two-rowed minor is zero, but some element is not zero. It is said to be of rank 0 if every element is zero.

In the discussion of the three equations (4), five cases arise:

(α) D of rank 3, *i.e.*, $D \neq 0$.

(β) D of rank 2 (*i.e.*, $D = 0$, but some two-rowed minor $\neq 0$), and

$$K_1 = \begin{vmatrix} k_1 & b_1 & c_1 \\ k_2 & b_2 & c_2 \\ k_3 & b_3 & c_3 \end{vmatrix}, \quad K_2 = \begin{vmatrix} a_1 & k_1 & c_1 \\ a_2 & k_2 & c_2 \\ a_3 & k_3 & c_3 \end{vmatrix}, \quad K_3 = \begin{vmatrix} a_1 & b_1 & k_1 \\ a_2 & b_2 & k_2 \\ a_3 & b_3 & k_3 \end{vmatrix}$$

not all zero.

(γ) D of rank 2 and K_1, K_2, K_3 all zero.

(δ) D of rank 1 (*i.e.*, every two-rowed minor $= 0$, but some element $\neq 0$), and

$$\begin{vmatrix} a_i & k_i \\ a_j & k_j \end{vmatrix}, \quad \begin{vmatrix} b_i & k_i \\ b_j & k_j \end{vmatrix}, \quad \begin{vmatrix} c_i & k_i \\ c_j & k_j \end{vmatrix} \quad (i,\ j \text{ chosen from } 1,\ 2,\ 3)$$

not all zero; there are nine such determinants K.

(ϵ) D of rank 1, and all nine of the determinants K zero.

In case (α) the equations have a single set of solutions (§ 23). In cases (β) and (δ) there is no set of solutions. In case (γ) one of the equations is a linear combination of the other two; for example, if $a_1b_2 - a_2b_1 \neq 0$, the first two equations determine x and y as linear functions of z (as shown by transposing the terms in z and solving the resulting equations for x and y), and the resulting values of x and y satisfy the third equation identically as to z. Finally, in case (ϵ), two of the equations are obtained by multiplying the remaining one by constants. For (β) the proof follows from (18). For (γ), (δ), (ϵ), the proof is given in § 25.

The reader acquainted with the elements of solid analytic geometry will see that the planes represented by the three equations have the following relations:

(α) The 3 planes intersect in a single point.

(β) Two of the planes intersect in a line parallel to the third plane.

(γ) The 3 planes intersect in a common line.

(δ) The 3 planes are parallel and not all coincident.

(ϵ) The 3 planes coincide.

25. Fundamental Theorem. *Let the determinant D of the coefficients of the unknowns in equations (17) be of rank r, $r < n$. If the determinants K obtained from the $(r + 1)$-rowed minors of D by replacing the elements of any column by the corresponding known terms k_i are not all zero, the equations are inconsistent. But if these determinants K are all zero, the r equations involving the elements of a non-vanishing r-rowed minor of D determine uniquely r of the variables as linear functions of the remaining $n - r$ variables, and the expressions for these r variables satisfy also the remaining $n - r$ equations.*

For example, let $r = n - 1$. Then $D = 0$ and the K's are the $K_1, \ldots,$ K_n of § 23. Hence, by (18), the equations are inconsistent unless $K_1, \ldots, K_n$ are all zero. This affords an illustration of the following

LEMMA 1. If every $(r + 1)$-rowed minor M formed from certain $r + 1$ rows of D is zero, the corresponding $r + 1$ equations (17) are inconsistent if there is a non-vanishing determinant K formed from any M by replacing the elements of any column by the corresponding known terms k_i.

For concreteness,* let the rows in question be the first $r + 1$ and let

$$K = \begin{vmatrix} a_{11} & \dots & a_{1r} & k_1 \\ \cdot & \cdot & \cdot & \cdot \\ a_{r+1\,1} & \dots & a_{r+1\,r} & k_{r+1} \end{vmatrix} \neq 0.$$

Let $d_1, \dots, d_{r+1}$ be the minors of $k_1, \dots, k_{r+1}$ in K. Multiply the first $r + 1$ equations (17) by $d_1, -d_2, \dots, (-1)^r d_{r+1}$, respectively, and add. The right member of the resulting equation is $\pm K$. The coefficient of x_s is

$$\pm \begin{vmatrix} a_{11} & \dots & a_{1r} & a_{1s} \\ \cdot & \cdot & \cdot & \cdot \\ a_{r+1\,1} & \dots & a_{r+1\,r} & a_{r+1\,s} \end{vmatrix}$$

and is zero, being an M. Hence $0 = \pm K$.

Lemma 2. If all of the determinants M and K in Lemma 1 are zero, but an r-rowed minor of an M is not zero, one of the corresponding $r + 1$ equations is a linear combination of the remaining r equations.

As before let the $r + 1$ rows in question be the first $r + 1$. Let the non-vanishing r-rowed minor be

$$d_{r+1} = \begin{vmatrix} a_{11} & \dots & a_{1r} \\ \cdot & \cdot & \cdot \\ a_{r1} & \dots & a_{rr} \end{vmatrix} \neq 0. \tag{19}$$

Let the functions obtained by transposing the terms k_i in (17) be

$$L_i \equiv a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n - k_i.$$

By the multiplication made in the proof of Lemma 1,

$$d_1L_1 - d_2L_2 + \cdots + (-1)^r d_{r+1}L_{r+1} = \mp K = 0.$$

Hence L_{r+1} is a linear combination of $L_1, \dots, L_r$.

The first part of the fundamental theorem is true by Lemma 1. The second part is readily proved by means of Lemma 2. Let (19) be the non-vanishing r-rowed minor of D. For $s > r$, the sth equation is a linear combination of the first r equations, and hence is satisfied by any set of solutions of the latter. In the latter transpose the terms involving $x_{r+1}, \dots, x_n$. Since the determinant of the coefficients of $x_1, \dots, x_r$ is not zero, § 23 shows that $x_1, \dots, x_r$ are uniquely determined linear functions of $x_{r+1}, \dots, x_n$ (which enter from the new right members).

* All other cases may be reduced to this one by rearranging the n equations and relabeling the unknowns (replacing x_3 by the new x_1, for example).

EXERCISES

1. Write out the proof of the theorem in § 25 for the cases (γ), (δ), (ϵ) in § 24.

Discuss the following systems of equations:

2. $$\begin{aligned} 2x + y + 3z &= 1, \\ 4x + 2y - z &= -3, \\ 2x + y - 4z &= -4. \end{aligned}$$

3. $$\begin{aligned} 2x + y + 3z &= 1, \\ 4x + 2y - z &= 3, \\ 2x + y - 4z &= 4. \end{aligned}$$

4. $$\begin{aligned} x - 3y + 4z &= 1, \\ 4x - 12y + 16z &= 3, \\ 3x - 9y + 12z &= 3. \end{aligned}$$

5. $$\begin{aligned} x - 3y + 4z &= 1, \\ 4x - 12y + 16z &= 4, \\ 3x - 9y + 12z &= 3. \end{aligned}$$

6. Discuss the equations in Exs. 4 and 5, p. 134, when two or more of the numbers a, b, c, k are equal.

7. Discuss the equations in Ex. 6, p. 134, when $a = -2$.

26. Homogeneous Linear Equations. When the known terms $k_1, \ldots, k_n$ in (17) are all zero, the equations are called *homogeneous*. The determinants K are now all zero, so that the n homogeneous equations are never inconsistent. This is also evident from the fact that they have the set of solutions $x_1 = 0, \ldots, x_n = 0$. By (18), there is no further set of solutions if $D \neq 0$. If $D = 0$, there are further sets of solutions: if D is of rank r, there occur $n - r$ arbitrary parameters in the general set of solutions (§ 25). A particular case of this result is the much used theorem:

A necessary and sufficient condition that n linear homogeneous equations in n unknowns shall have a set of solutions, other than the trivial one in which each unknown is zero, is that the determinant of the coefficients be zero.

27. The case of a system of fewer than n linear equations in n unknowns may be treated by means of the Lemmas in § 25.

In case we have a system of more than n linear equations in n unknowns, we may first discuss n of the equations. If these are inconsistent, the entire system is. If they are consistent, the general set S of solutions may be found and substituted into the remaining equations. There result conditions on the parameters occurring in S, and these linear conditions may be treated in the usual manner. Ultimately we get either the general set of solutions of the entire system of equations or find that they are inconsistent. To decide in advance which of these cases will arise we have only to find the maximum order r of a non-vanishing r-rowed determinant formed from the coefficients of the unknowns, taken in the regular order

in which they occur in the equations, and ascertain whether or not the $(r+1)$-rowed determinants K, formed as in § 25, are all zero.*

28. An important case is that of n non-homogeneous linear equations in $n-1$ unknowns $x_1, \ldots, x_{n-1}$. By multiplying the known terms by $x_n = 1$, we bring this case under that of n homogeneous linear equations in n unknowns (§ 26). Then (18) gives $Dx_n = 0$, $D = 0$, so that the given equations are inconsistent if $D \neq 0$.

There is no set of solutions of the n equations

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1\,n-1}x_{n-1} = k_1, \\ \cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot \\ a_{n\,1}x_1 + a_{n\,2}x_2 + \cdots + a_{n\;n-1}x_{n-1} = k_n, \end{array} \qquad \begin{vmatrix} a_{11} & \ldots & a_{1\,n-1} & k_1 \\ \cdot & \cdot & \cdot & \cdot \\ a_{n\,1} & \ldots & a_{n\;n-1} & k_n \end{vmatrix} \neq 0.$$

EXERCISES

Discuss the following systems of equations:

1. $x + y + 3z = 0,\quad x + 2y + 2z = 0,\quad x + 5y - z = 0.$

2. $2x - y + 4z = 0,\quad x + 3y - 2z = 0,\quad x - 11y + 14z = 0.$

3. $x - 3y + 4z = 0,\quad 4x - 12y + 16z = 0,\quad 3x - 9y + 12z = 0.$

4. $6x + 4y + 3z - 84w = 0,\quad x + 2y + 3z - 48w = 0,\quad x - 2y + z - 12w = 0,\quad 4x + 4y - z - 24w = 0.$

5. $2x + 3y - 4z + 5w = 0,\quad 3x + 5y - z + 2w = 0,\quad 7x + 11y - 9z + 12w = 0,\quad 3x + 4y - 11z + 13w = 0.$

6. $2x + y + 3z = 1,\quad 4x + 2y - z = -3,\quad 2x + y - 4z = -4,\quad 10x + 5y - 6z = -10.$

7. $2x - y + 3z = 2,\quad x + 7y + z = 1,\quad 3x + 5y - 5z = -3,\quad 4x - 3y + 2z = 1.$

8. Obtain a consistent system of equations from the system in Ex. 7 by replacing the term -3 by a new value.

9. In three linear homogeneous equations in x, y, z, w, the latter are proportional to four determinants of order 3 formed from the coefficients.

* For an abbreviated statement, the concepts matrix and its rank are needed. *Cf.* Bôcher, *Introduction to Higher Algebra*, p. 46.

CHAPTER XII

Resultants and Discriminants

1. Introduction. If the two equations

$$ax + b = 0, \qquad cx + d = 0 \qquad\qquad (a \neq 0,\ c \neq 0)$$

are simultaneous, *i.e.*, if x has the same value in each, then

$$x = -\frac{b}{a} = -\frac{d}{c}, \qquad\qquad R \equiv ad - bc = 0,$$

and conversely. Hence a necessary and sufficient condition that the equations have a common root is $R = 0$. We call R the *resultant* (or *eliminant*) of the two equations.

The result of eliminating x between the two equations might equally well have been written in the form $bc - ad = 0$. But the arbitrary selection of R as the resultant, rather than the product of R by some constant as -1, is a matter of more importance than apparent at first sight. We seek a *definite* function of the coefficients a, b, c, d of the *functions* $ax + b$, $cx + d$, and not merely a property $R = 0$ or $R \neq 0$ of the corresponding *equations*. Accordingly, we shall lay down the definition in § 2, which, as the reader may verify, leads to R in our present example.

Methods of elimination which seem plausible often yield not R itself, but the product of R by an extraneous function of the coefficients. This point (illustrated in Ex. 3, p. 156) indicates that the subject demands a more careful treatment than is often given.

We may even introduce an extraneous factor zero. Let $\alpha \neq 0$,

$$f(x) = x^2 - 2\,\alpha x - 3\,\alpha^2, \qquad g(x) = x - \alpha.$$

From f subtract $(x + \alpha)g$. Multiply the remainder, $-2\,\alpha(x + \alpha)$, by $x - 3\,\alpha$ and add the product to $2\,\alpha f$. The sum is zero. But the resultant is $-4\,\alpha^2$ (the value of f for $x = \alpha$) and is not zero. As we used g only in the first step and there, in effect, replaced it by $x^2 - \alpha^2$, we really found the resultant of the latter and f. The extraneous factor introduced (*cf.* Ex. 7, p. 152) is the resultant of $x + \alpha$ and f and this resultant is zero.

2. Resultant of Two Polynomials in x. Let

$$(1)\qquad \begin{cases} f(x) = a_0x^m + a_1x^{m-1} + \cdots + a_m & (a_0 \neq 0) \\ g(x) = b_0x^n + b_1x^{n-1} + \cdots + b_n & (b_0 \neq 0) \end{cases}$$

be two polynomials of degrees m and n. Let $\alpha_1, \ldots, \alpha_m$ be the roots of $f(x) = 0$. Now α_1 is a root of $g(x) = 0$ only when $g(\alpha_1) = 0$. The two equations have a root in common if and only if the product

$$g(\alpha_1)\, g(\alpha_2) \ \ldots \ g(\alpha_m)$$

is zero. This symmetric function of the roots of $f(x) = 0$ is of degree n in any one root and hence is expressible as a polynomial of degree n in the elementary symmetric functions (Chap. VII, § 3), which equal $-a_1/a_0$, $a_2/a_0, \ldots$. To be rid of the denominators a_0, it suffices to multiply our polynomial by a_0^n. We therefore define

$$(2)\qquad R(f, g) = a_0^n g(\alpha_1)g(\alpha_2) \ \ldots \ g(\alpha_m)$$

to be the resultant of f and g. It equals a rational integral function of $a_0, \ldots, a_m, b_0, \ldots, b_n$.

EXERCISES

1. If $m = 1, n = 2,\quad R(f, g) = b_0a_1^2 - b_1a_0a_1 + b_2a_0^2$.

2. If $m = 2, n = 1,\quad R(f, g) = a_0(b_0\alpha_1 + b_1)(b_0\alpha_2 + b_1) = a_0b_1^2 - a_1b_0b_1 + a_2b_0^2$,

since
$$a_0(\alpha_1 + \alpha_2) = -a_1, \qquad a_0\alpha_1\alpha_2 = a_2.$$

3. If $\beta_1, \ldots, \beta_n$ are the roots of $g(x) = 0$, so that

$$g(\alpha_i) = b_0(\alpha_i - \beta_1)(\alpha_i - \beta_2) \ \ldots \ (\alpha_i - \beta_n),$$

then

$$\begin{aligned} R(f, g) = a_0^n b_0^m\, &(\alpha_1 - \beta_1)(\alpha_1 - \beta_2) \ \ldots \ (\alpha_1 - \beta_n) \\ &\cdot (\alpha_2 - \beta_1)(\alpha_2 - \beta_2) \ \ldots \ (\alpha_2 - \beta_n) \\ &\cdots\cdots\cdots\cdots\cdots \\ &\cdot (\alpha_m - \beta_1)(\alpha_m - \beta_2) \ \ldots \ (\alpha_m - \beta_n). \end{aligned}$$

Multiplying together the differences in each column, we see that

$$R(f, g) = (-1)^{mn} b_0^m f(\beta_1) f(\beta_2) \ \ldots \ f(\beta_n) = (-1)^{mn} R(g, f).$$

4. If $m = 2$, $n = 1$, $R(g, f) = b_0^2 f(-b_1/b_0) = a_0b_1^2 - a_1b_0b_1 + a_2b_0^2$, which equals $R(f, g)$ by Ex. 2. This illustrates the final result in Ex. 3.

5. If $m = n = 2$,
$$\begin{aligned} R(f, g) &= a_0^2b_0^2\alpha_1^2\alpha_2^2 + a_0^2b_0b_1\alpha_1\alpha_2(\alpha_1 + \alpha_2) \\ &\quad + a_0^2b_0b_2(\alpha_1^2 + \alpha_2^2) + a_0^2b_1^2\alpha_1\alpha_2 + a_0^2b_1b_2(\alpha_1 + \alpha_2) + a_0^2b_2^2 \\ &= b_0^2a_2^2 - b_0b_1a_1a_2 + b_0b_2(a_1^2 - 2\,a_0a_2) + b_1^2a_0a_2 - b_1b_2a_0a_1 + a_0^2b_2^2. \end{aligned}$$

This equals $R(g, f)$, since it is unaltered when the a's and b's are interchanged.

6. R is homogeneous and of degree n in $a_0, \ldots, a_m$; homogeneous and of degree m in $b_0, \ldots, b_n$. R has the terms

$$a_0^n b_n^m, \qquad (-1)^{mn} b_0^m a_m^n.$$

7. $R(f, g_1 g_2) = R(f, g_1) \cdot R(f, g_2)$.
8. $R(f, x^n) = (-1)^{mn} a_m^n$.

3. Irreducibility of the Resultant of Two Polynomials in One Variable.* The resultant of two polynomials $f(x)$ and $g(x)$ was seen (§ 2) to equal a polynomial $r(a_0, \ldots, a_m, b_0, \ldots, b_n)$ in the coefficients of f and g. Let these coefficients be regarded as independent variables. Then r is *irreducible, i.e.,* is not equal to the product of two polynomials r_1 and r_2 in $a_0, \ldots, b_n$ with numerical coefficients, if neither r_1 nor r_2 is a numerical constant.** Suppose that $r \equiv r_1 r_2$. Since r is homogeneous in $a_0, \ldots, a_m$, each factor r_i is. Likewise, each r_i is homogeneous in $b_0, \ldots, b_n$. Hence

$$r\left(1, \frac{a_1}{a_0}, \ldots, \frac{a_m}{a_0}, 1, \frac{b_1}{b_0}, \ldots, \frac{b_n}{b_0}\right) \equiv r_1\left(1, \frac{a_1}{a_0}, \ldots\right) \cdot r_2\left(1, \frac{a_1}{a_0}, \ldots\right).$$

Replace $a_1/a_0, \ldots, a_m/a_0$ by the corresponding symmetric functions of the roots $\alpha_1, \ldots, \alpha_m$, also $b_1/b_0, \ldots, b_n/b_0$ by the corresponding symmetric functions of $\beta_1, \ldots, \beta_n$. Let the factors on the right become the polynomials P_1 and P_2 in $\alpha_1, \ldots, \beta_n$. Then (Ex. 3),

$$(\alpha_1 - \beta_1) \ldots (\alpha_1 - \beta_n)(\alpha_2 - \beta_1) \ldots (\alpha_m - \beta_n) \equiv P_1 P_2,$$

identically in the α's and β's. Apart from numerical factors, P_1 is therefore the product of certain of the differences $\alpha_1 - \beta_1, \ldots$, and P_2 the product of the others. But this is impossible since P_1 is symmetric in $\alpha_1, \ldots, \alpha_m$ and symmetric in $\beta_1, \ldots, \beta_n$.

4. A Correct Conclusion to be Drawn from Any Method of Elimination. Since the determination of r by means of symmetric functions of the roots is excessively laborious unless m or n is very small, we shall later give other methods. But we shall not know, without a careful enquiry, whether or not such a new method introduces an extraneous factor. Each

* In place of §§ 3, 4, the reader may use § 9. But this substitution should be made only if the briefest course is desired.

** This is evident for the resultant $ad - bc$ in § 1. For, if it were the product of two linear functions, the one not involving a would necessarily be d (or a numerical constant times d) and similarly the other factor would then be a.

method leads in fact to a polynomial $F(a_0, \ldots, b_n)$ with the property that every set of solutions $a_0, \ldots, b_n$ of $r = 0$ is a set of solutions of $F = 0$. It then follows that *r is a factor of F.*

For example, if $R(f, g) = 0$,

$$f \equiv a_0x^2 + a_1x + a_2 = 0, \qquad g \equiv b_0x^2 + b_1x + b_2 = 0$$

have a common root x. Then

$$b_2f - a_2g = (a_0b_2 - a_2b_0)x^2 + (a_1b_2 - a_2b_1)x = 0,$$
$$-b_0f + a_0g = (a_0b_1 - a_1b_0)x + a_0b_2 - a_2b_0 = 0.$$

Exclude for the moment the case $a_2 = b_2 = 0$. Then $x \neq 0$ and

$$(3) \qquad F \equiv \begin{vmatrix} a_0b_2 - a_2b_0 & a_1b_2 - a_2b_1 \\ a_0b_1 - a_1b_0 & a_0b_2 - a_2b_0 \end{vmatrix} = 0.$$

It is easily verified that $F = 0$ also in the excluded case. Hence any set of solutions $a_0, \ldots, b_2$ of $r = 0$ is a set of solutions of $F = 0$. We found r in Ex. 5 above. It is seen to be identical with this F.

To prove in general that r is a factor of F, set

$$r = c_0a_0^n + c_1a_0^{n-1} + \cdots + c_n,$$

where $c_0, \ldots, c_n$ are polynomials in $a_1, \ldots, b_n$, while c_0 is not identically zero (Ex. 6 above). Express also F as a polynomial in a_0 and apply the greatest common divisor process to F and r. Suppose that r is not a factor of F. If * the degree of F in a_0 is $\geqq n$, we may write

$$k_0F = q_0r + r_1, \qquad k_1r = q_1r_1 + r_2, \qquad k_2r_1 = q_2r_2 + r_3,$$

where q_0, q_1, q_2, r_1, r_2 may involve a_0, while k_0, k_1, k_2, r_3 do not (for simplicity we assume that r_3 is the first r_i not involving a_0). If r_3 were identically zero, r_2 (or a factor actually involving a_0) would be a factor of r, as shown by the last two of our three equations. Since r_2 is of lower degree in a_0 than r, this contradicts the irreducibility of r (§ 3). Hence there exist constants $a_1', \ldots, b_n'$ such that

$$r_3(a_1', \ldots, b_n') \neq 0, \qquad c_0(a_1', \ldots, b_n') \neq 0.$$

For $a_1 = a_1', \ldots, b_n = b_n'$, r becomes a polynomial in a_0 with constant coefficients and hence (Ch. V) vanishes for some value a_0' of a_0. By

* In the contrary case, we drop the first equation and set $r_1 \equiv F$.

hypothesis, any set of solutions, as $a_0', a_1', \ldots, b_n'$ of $r = 0$ is a set of solutions of $F = 0$. Hence $F(a_0', \ldots, b_n') = 0$. For these values $a_0', \ldots, b_n'$ of $a_0, \ldots, b_n$, we have $r_1 = 0$ by the first of our three equations, then $r_2 = 0$ by the second, and $r_3 = 0$ by the third. The last result contradicts $r_3(a_1', \ldots, b_n') \neq 0$.

If any method of eliminating x between two equations in x leads to a relation $F = 0$, where F is a polynomial in the coefficients, then F has as a factor the true resultant of the equations.

5. Sylvester's Dialytic Method of Elimination. Let the equations

$$a_0x^3 + a_1x^2 + a_2x + a_3 = 0, \qquad b_0x^2 + b_1x + b_2 = 0$$

have a common root x, so that their resultant r is zero.

Multiply the first equation by x and the second by x^2 and x in turn. We now have five equations

$$\begin{aligned} a_0x^4 + a_1x^3 + a_2x^2 + a_3x \phantom{{}+a_3} &= 0,\\ a_0x^3 + a_1x^2 + a_2x + a_3 &= 0,\\ b_0x^4 + b_1x^3 + b_2x^2 \phantom{{}+a_3x+a_3} &= 0,\\ b_0x^3 + b_1x^2 + b_2x \phantom{{}+a_3} &= 0,\\ b_0x^2 + b_1x + b_2 &= 0, \end{aligned}$$

which are linear and homogeneous in $x^4, x^3, x^2, x, 1$. Hence

$$F = \begin{vmatrix} a_0 & a_1 & a_2 & a_3 & 0 \\ 0 & a_0 & a_1 & a_2 & a_3 \\ b_0 & b_1 & b_2 & 0 & 0 \\ 0 & b_0 & b_1 & b_2 & 0 \\ 0 & 0 & b_0 & b_1 & b_2 \end{vmatrix} \tag{4}$$

is zero. By § 4, r is a factor of F. But the diagonal term $a_0^2b_2^3$ of F is a term of r (Ex. 6, p. 152). Hence F is the resultant.

In general, if the equations are

$$a_0x^m + \cdots + a_m = 0, \quad b_0x^n + \cdots + b_n = 0,$$

we multiply the first equation by $x^{n-1}, x^{n-2}, \ldots, x, 1$, in turn, and the second by $x^{m-1}, x^{m-2}, \ldots, x, 1$, in turn. We obtain $n + m$ equations which

are linear and homogeneous in the $m + n$ quantities $x^{m+n-1}, \ldots, x, 1$. Hence the determinant

$$
(5) \qquad F = \left|\begin{matrix}
a_0 & a_1 & a_2 & \ldots & a_m & 0 & \ldots & \ldots & 0 \\
0 & a_0 & a_1 & a_2 & \ldots & a_m & 0 & \ldots & 0 \\
0 & 0 & a_0 & a_1 & a_2 & \ldots & a_m & 0 \ldots & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \ldots & 0 & a_0 & a_1 & a_2 & \ldots & \ldots & a_m \\
b_0 & b_1 & \ldots & \ldots & b_n & 0 & \ldots & \ldots & 0 \\
0 & b_0 & b_1 & \ldots & \ldots & b_n & \ldots & \ldots & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \ldots & 0 & b_0 & b_1 & \ldots & \ldots & \ldots & b_n
\end{matrix}\right|
\begin{matrix}
\left.\vphantom{\begin{matrix}a\\a\\a\\a\\a\end{matrix}}\right\} n \text{ rows} \\
\left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} m \text{ rows}
\end{matrix}
$$

is zero. By § 4, r is a factor of F. But the diagonal term $a_0{}^n b_n{}^m$ is a term of r. Hence F is the resultant.

EXERCISES

1. For $m = n = 2$, the resultant is

$$
r = \begin{vmatrix}
a_0 & a_1 & a_2 & 0 \\
0 & a_0 & a_1 & a_2 \\
b_0 & b_1 & b_2 & 0 \\
0 & b_0 & b_1 & b_2
\end{vmatrix}.
$$

Interchange the second and third rows, apply Laplace's development, and prove that

$$
r = (a_0b_2)^2 - (a_0b_1)(a_1b_2),
$$

where (a_0b_2) denotes $a_0b_2 - a_2b_0$, etc. Compare with (3).

2. For $m = n = 3$, show by interchanges of rows that

$$
r = - \begin{vmatrix}
a_0 & a_1 & a_2 & a_3 & 0 & 0 \\
b_0 & b_1 & b_2 & b_3 & 0 & 0 \\
0 & a_0 & a_1 & a_2 & a_3 & 0 \\
0 & b_0 & b_1 & b_2 & b_3 & 0 \\
0 & 0 & a_0 & a_1 & a_2 & a_3 \\
0 & 0 & b_0 & b_1 & b_2 & b_3
\end{vmatrix}.
$$

Apply Laplace's development, selecting minors from the first two rows, and to the complementary minors apply a similar development. This may be done by inspection and the following value of $-r$ be obtained:

$$\begin{aligned}&(a_0b_1)\{(a_1b_2)(a_2b_3) - (a_1b_3)^2 + (a_2b_3)(a_0b_3)\}\\ -&(a_0b_2)\{(a_0b_2)(a_2b_3) - (a_0b_3)(a_1b_3)\}\\ +&(a_0b_3)\{(a_0b_1)(a_2b_3) - (a_0b_3)^2\}.\end{aligned}$$

The third term of the first line and the first term of the last line are alike. Hence, changing the signs,

$$\begin{aligned}r = {}&(a_0b_3)^3 - 2\,(a_0b_1)(a_0b_3)(a_2b_3) - (a_0b_2)(a_0b_3)(a_1b_3)\\ &+ (a_0b_2)^2\,(a_2b_3) + (a_0b_1)(a_1b_3)^2 - (a_0b_1)(a_1b_2)(a_2b_3).\end{aligned}$$

3. For $m = n = 3$, the method which led to (3) gives

$$- b_0f + a_0g = (a_0b_1)\,x^2 + (a_0b_2)\,x + (a_0b_3),$$
$$(b_3f - a_3g)/x = (a_0b_3)\,x^2 + (a_1b_3)\,x + (a_2b_3).$$

By (3), the resultant of these two quadratic functions is

$$F = \begin{vmatrix} (a_0b_3) & (a_0b_1) \\ (a_2b_3) & (a_0b_3) \end{vmatrix}^2 - \begin{vmatrix} (a_0b_3) & (a_0b_1) \\ (a_1b_3) & (a_0b_2) \end{vmatrix} \cdot \begin{vmatrix} (a_1b_3) & (a_0b_2) \\ (a_2b_3) & (a_0b_3) \end{vmatrix}.$$

This is, however, not the resultant r of the cubic functions f, g. To show that (a_0b_3) is an extraneous factor, note that the terms of F not having this factor explicitly are

$$(a_0b_1)(a_2b_3)\{(a_0b_1)(a_2b_3) - (a_0b_2)(a_1b_3)\}.$$

The quantity in brackets equals $-(a_0b_3)(a_1b_2)$, since

$$0 = \tfrac{1}{2}\begin{vmatrix} a_0 & a_1 & a_2 & a_3 \\ b_0 & b_1 & b_2 & b_3 \\ a_0 & a_1 & a_2 & a_3 \\ b_0 & b_1 & b_2 & b_3 \end{vmatrix} = (a_0b_1)(a_2b_3) - (a_0b_2)(a_1b_3) + (a_0b_3)(a_1b_2).$$

We now see that $F = r \cdot (a_0b_3)$, where r is given in Ex. 2.

4. Verify that (a_0b_3) is an extraneous factor by showing that if $x^3 - 1 = 0$, $x^3 - x = 0$, then $r = 0$, $(a_0b_3) \neq 0$.

5. The resultant of $L \equiv \alpha x + \beta y$ and $L' \equiv \alpha' x + \beta' y$ is $R = \alpha\beta' - \alpha'\beta$. The determinant of the coefficients of x^2, xy, y^2 in $L^2 = 0$, $LL' = 0$, $L'^2 = 0$ is

$$R' = \begin{vmatrix} \alpha^2 & 2\,\alpha\beta & \beta^2 \\ \alpha\alpha' & \alpha\beta' + \alpha'\beta & \beta\beta' \\ \alpha'^2 & 2\,\alpha'\beta' & \beta'^2 \end{vmatrix}.$$

If $R = 0$ there exist values not both zero of x and y such that $L = L' = 0$ and hence values of x^2, xy, y^2, not all zero, such that $L^2 = 0$, etc. Thus $R = 0$ implies $R' = 0$. Since R is irreducible, it is a factor of R'. But if $R' = 0$, we are not to infer hastily that the values of x^2, xy, y^2 obtained from the three equations linear in them are consistent (*i.e.*, the product of the first and third equals the square of xy) and hence have no right to conclude that $R' = 0$ implies $R = 0$ and thus that R' is a power of R (as done in some texts).

If $R' = 0$, the three linear homogeneous equations whose coefficients are the elements in the three rows of the determinant R' have solutions not all zero, which may be designated x^2, $xy - z$, y^2. Then the equations may be written in the form

$$L^2 = 2\,\alpha\beta z, \quad LL' = (\alpha\beta' + \alpha'\beta)z, \quad L'^2 = 2\,\alpha'\beta' z.$$

Thus

$$0 = (LL')^2 - L^2L'^2 = R^2z^2.$$

If $R \neq 0$, then $z = 0$, $L = L' = 0$, $Rx = Ry = 0$, whereas x, y, z are not all zero. Hence $R' = 0$ implies $R = 0$. Thus each irreducible factor of R' is a numerical multiple of R. By examining one term of R', we see that $R' = R^3$.

6. The determinant of the coefficients of x^3, x^2y, xy^2, y^3 in

$$L^3 = 0, \qquad L^2L' = 0, \qquad LL'^2 = 0, \qquad L'^3 = 0,$$

equals R^6. Prove as in Ex. 5 and also as in Ex. 7.

7. Reduce the determinant R' in Ex. 5 to the form R^3. If $\beta = 0$, R' is evidently R^3. If $\beta \neq 0$, multiply the elements of the second column by $-\alpha/\beta$, those of the third column by α^2/β^2, and add the products to the elements of the first column. The elements of the new first column are 0, 0, R^2/β^2. Hence

$$R' = \frac{R^2}{\beta}\begin{vmatrix} 2\,\alpha\beta & \beta \\ \alpha\beta' + \alpha'\beta & \beta' \end{vmatrix} = \frac{R^2}{\beta}\begin{vmatrix} \alpha\beta & \beta \\ \alpha'\beta & \beta' \end{vmatrix} = R^2 \cdot R.$$

8. If for $F = 0$ we omit one of the equations in § 5, we have a consistent set of equations which determine x in general. Thus if $m = n = 2$, $xf = 0$, $f = 0$, $g = 0$ give $a_0(a_0b_1)x = -a_0(a_0b_2)$. The latter is in agreement with the linear equation in the example, p. 153.

6. Discriminants. Let $\alpha_1, \ldots, \alpha_m$ be the roots of

$$f(x) \equiv a_0x^m + a_1x^{m-1} + \cdots + a_m = 0 \qquad (a_0 \neq 0) \tag{6}$$

As in Ch. III, § 3, we define the discriminant of (6) to be

$$D = a_0^{2m-2}(\alpha_1 - \alpha_2)^2(\alpha_1 - \alpha_3)^2 \ldots (\alpha_1 - \alpha_m)^2(\alpha_2 - \alpha_3)^2 \ldots (\alpha_{m-1} - \alpha_m)^2.$$

Evidently D is unaltered by the interchange of any two roots. Since the degree in any root is $2(m-1)$, the symmetric function D equals a polynomial in $a_0, \ldots, a_m$. Indeed, a_0^{2m-2} is the lowest power of a_0 sufficient

to cancel the denominators introduced by replacing $\Sigma\alpha_1$ by $-a_1/a_0, \ldots,$ $\alpha_1\alpha_2 \ldots \alpha_m$ by $\pm a_m/a_0$. Now *

$$f'(\alpha_1) = a_0(\alpha_1 - \alpha_2)(\alpha_1 - \alpha_3) \ldots (\alpha_1 - \alpha_m),$$
$$f'(\alpha_2) = a_0(\alpha_2 - \alpha_1)(\alpha_2 - \alpha_3) \ldots (\alpha_2 - \alpha_m),$$
$$f'(\alpha_3) = a_0(\alpha_3 - \alpha_1)(\alpha_3 - \alpha_2)(\alpha_3 - \alpha_4) \ldots (\alpha_3 - \alpha_m), \ldots .$$

Hence

$$a_0^{m-1}f'(\alpha_1) \ldots f'(\alpha_m) = a_0^{2m-1}(-1)^{1+2+\cdots+m-1}(\alpha_1-\alpha_2)^2 \ldots (\alpha_{m-1}-\alpha_m)^2$$
$$= (-1)^{\frac{m(m-1)}{2}} a_0 D.$$

By (2), the left member is the resultant of $f(x), f'(x)$. Hence

$$D = (-1)^{\frac{m(m-1)}{2}} \frac{1}{a_0} R(f, f'). \tag{7}$$

For another proof that D is a numerical multiple of R/a_0, see Ex. 9 below.

EXERCISES

1. Show that the discriminant of $f \equiv y^3 + py + q = 0$ is $-4p^3 - 27q^2$ by evaluating the determinant of order five for $R(f, f')$.

2. Find the relation between the discriminant of $f(x) = 0$ and the resultant of $mf(x) - xf'(x)$ and $f'(x)$.

3. Hence the discriminant of $a_0x^3 + a_1x^2 + a_2x + a_3$ is $-\frac{1}{3}r$, where

$$r = (a_1a_2 - 9a_0a_3)^2 - (2a_2^2 - 6a_1a_3)(2a_1^2 - 6a_0a_2)$$

is the resultant of $a_1x^2 + 2a_2x + 3a_3 = 0$, $3a_0x^2 + 2a_1x + a_2 = 0$, by (3).

4. The discriminant of the product of two functions equals the product of their discriminants multiplied by the square of their resultant. Hint: use the expressions in terms of the differences of the roots.

5. a_0 is a factor of $R(f, f')$ by the first column of its determinant.

6. For $a_0 = 1$, the discriminant equals

$$\begin{vmatrix} 1 & \alpha_1 & \alpha_1^2 & \ldots & \alpha_1^{m-1} \\ 1 & \alpha_2 & \alpha_2^2 & \ldots & \alpha_2^{m-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & \alpha_m & \alpha_m^2 & \ldots & \alpha_m^{m-1} \end{vmatrix}^2 = \begin{vmatrix} s_0 & s_1 & s_2 & \ldots & s_{m-1} \\ s_1 & s_2 & s_3 & \ldots & s_m \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ s_{m-1} & s_m & s_{m+1} & \ldots & s_{2m-2} \end{vmatrix},$$

where $s_i = \alpha_1^i + \cdots + \alpha_m^i$. See Exs. 10, 11, p. 141.

* By differentiating $f(x) = a_0(x - \alpha_1) \ldots (x - \alpha_m)$ or by the first part of § 5, Ch. VII.

7. Hence the discriminant of $x^3 + px + q = 0$ equals

$$\begin{vmatrix} 3 & 0 & -2p \\ 0 & -2p & -3q \\ -2p & -3q & 2p^2 \end{vmatrix} = -4p^3 - 27q^2.$$

8. The discriminant $D(a_0, \ldots, a_m)$ is irreducible. As in § 3, a factor would equal a product P of powers of the differences $\alpha_i - \alpha_j$ such that P is symmetric in $\alpha_1, \ldots, \alpha_m$. Thus every difference would be a factor. But the product of the first powers of all the differences is changed in sign by any interchange of two roots (Ch. XI, § 12). Hence P is divisible by the square of the last product.

9. Prove that D is a constant times $R(f, f') \div a_0$ by use of § 4. Since $D = 0$ implies $R = 0$, the irreducible D is a factor of R. But D is of total degree $2m - 2$ in $a_0, a_1, \ldots$, and R is of total degree $2m - 1$. Hence R/D is of the first degree and thus (Ex. 5) a numerical multiple of a_0.

7.† Euler's Method of Elimination. Let f and g be given by (1). If $f(x) = 0$ and $g(x) = 0$ have a common root c, then

$$f(x) \equiv (x - c) f_1(x), \qquad g(x) \equiv (x - c) g_1(x),$$

identically in x, where $f_1(x)$ is a polynomial of degree $m - 1$, and $g_1(x)$ is of degree $n - 1$. Hence

$$f(x)g_1(x) \equiv g(x)f_1(x),$$

identically in x. Hence if $a_0, \ldots, b_n$ are any numbers for which $R(f, g) = 0$, there exist constants $q_1, \ldots, q_n, p_1, \ldots, p_m$ not all zero for which

$$\begin{aligned} &(a_0x^m + a_1x^{m-1} + \cdots + a_m)(q_1x^{n-1} + q_2x^{n-2} + \cdots + q_n) \\ \equiv\ &(b_0x^n + b_1x^{n-1} + \cdots + b_n)(p_1x^{m-1} + p_2x^{m-2} + \cdots + p_m), \end{aligned}$$

identically in x. Equating the coefficients of like powers of x in the two products, we obtain the relations

$$\begin{array}{lllr} a_0q_1 & - b_0p_1 & & = 0, \\ a_1q_1 + a_0q_2 & - b_1p_1 - b_0p_2 & & = 0, \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots \\ \quad a_mq_{n-1} + a_{m-1}q_n & & - b_np_{m-1} - b_{n-1}p_m & = 0, \\ \qquad\quad a_mq_n & & - b_np_m & = 0. \end{array}$$

Since these $m+n$ linear homogeneous equations in the unknowns $q_1, \ldots, q_n, -p_1, \ldots, -p_m$ have a set of solutions not all zero, the determinant of the coefficients is zero. Interchanging the rows and columns of this determinant, we get (5). The proof that (5) is the resultant follows as in the last two lines of § 5.

8.† Bézout's Method of Elimination. When the two equations are of the same degree, the method will be clear from the example

$$f \equiv a_0x^3 + a_1x^2 + a_2x + a_3 = 0, \qquad g \equiv b_0x^3 + b_1x^2 + b_2x + b_3 = 0.$$

Then

$$\begin{array}{c} a_0g - b_0f, \\ (a_0x + a_1)g - (b_0x + b_1)f, \\ (a_0x^2 + a_1x + a_2)g - (b_0x^2 + b_1x + b_2)f \end{array} \tag{8}$$

equal respectively

$$\begin{array}{l} (a_0b_1)x^2 \qquad\qquad\qquad + (a_0b_2)x + (a_0b_3) = 0, \\ (a_0b_2)x^2 + \{(a_0b_3) + (a_1b_2)\}x + (a_1b_3) = 0, \\ (a_0b_3)x^2 \qquad\qquad\qquad + (a_1b_3)x + (a_2b_3) = 0, \end{array}$$

where $(a_0b_1) = a_0b_1 - a_1b_0$, etc. The determinant of the coefficients is the negative of the resultant $R(f, g)$. Indeed, it is divisible by R (§ 4) and has a term of $-R$. The negative of the determinant is seen to have the expansion given as r in Ex. 2, p. 156.

The three equations used above are evident combinations of

$$x^2f = 0, \quad xf = 0, \quad f = 0, \quad x^2g = 0, \quad xg = 0, \quad g = 0,$$

the latter being the equations used in Sylvester's method of elimination. The determinant of the coefficients in these six equations is

$$R = \begin{vmatrix} a_0 & a_1 & a_2 & a_3 & 0 & 0 \\ 0 & a_0 & a_1 & a_2 & a_3 & 0 \\ 0 & 0 & a_0 & a_1 & a_2 & a_3 \\ b_0 & b_1 & b_2 & b_3 & 0 & 0 \\ 0 & b_0 & b_1 & b_2 & b_3 & 0 \\ 0 & 0 & b_0 & b_1 & b_2 & b_3 \end{vmatrix}.$$

The operations carried out to obtain the above three quadratic equations are seen to be step for step the following operations on determinants. First, a_0^3R is derived from the determinant R by multiplying the elements of the last three rows by a_0. To the elements of the new fourth row add the products of the elements of the 1st, 2nd, 3rd, 5th, 6th rows by $-b_0$, $-b_1$, $-b_2$, a_1, a_2 respectively [corresponding to the formation of the third function (8)]. To the elements of the fifth row add the products of the elements of the 2nd, 3rd, 6th rows by $-b_0$, b_1, a_1 respectively [corresponding to the second function (8)]. Finally, to the elements of the sixth row add the products of the elements of the third row by $-b_0$ [corresponding to $a_0g - b_0f$]. Hence

$$a_0^3 R = \begin{vmatrix} a_0 & a_1 & a_2 & a_3 & 0 & 0 \\ 0 & a_0 & a_1 & a_2 & a_3 & 0 \\ 0 & 0 & a_0 & a_1 & a_2 & a_3 \\ 0 & 0 & 0 & (a_0b_3) & (a_1b_3) & (a_2b_3) \\ 0 & 0 & 0 & (a_0b_2) & (a_0b_3)+(a_1b_2) & (a_1b_3) \\ 0 & 0 & 0 & (a_0b_1) & (a_0b_2) & (a_0b_3) \end{vmatrix},$$

so that R equals the 3-rowed minor enclosed by the dots. The method of Bézout therefore suggests a definite process for the reduction of Sylvester's determinant of order $2n$ (when $m = n$) to one of order n.

Next, for equations of different degrees, consider the example

$$f = a_0x^4 + a_1x^3 + a_2x^2 + a_3x + a_4, \qquad g = b_0x^2 + b_1x + b_2.$$

Then

$$a_0x^2g - b_0f, \qquad (a_0x + a_1)x^2g - (b_0x + b_1)f$$

equal respectively

$$(a_0b_1)\,x^3 + (a_0b_2)\,x^2 - a_3b_0x - a_4b_0,$$

$$(a_0b_2)\,x^3 + \{(a_1b_2) - a_3b_0\}\,x^2 - \{a_3b_1 + a_4b_0\}\,x - a_4b_1.$$

The determinant of the coefficients of x^3, x^2, x, 1 in these and xg, g, after the first and second rows are interchanged, is the determinant of order 4 enclosed by dots in the second determinant below. It is the resultant $R(f, g)$ by § 4.

As in the former example, we shall indicate the corresponding operations on Sylvester's determinant

$$R = \begin{vmatrix} a_0 & a_1 & a_2 & a_3 & a_4 & 0 \\ 0 & a_0 & a_1 & a_2 & a_3 & a_4 \\ b_0 & b_1 & b_2 & 0 & 0 & 0 \\ 0 & b_0 & b_1 & b_2 & 0 & 0 \\ 0 & 0 & b_0 & b_1 & b_2 & 0 \\ 0 & 0 & 0 & b_0 & b_1 & b_2 \end{vmatrix}.$$

Multiply the elements of the third and fourth rows by a_0. In the resulting determinant a_0^2R, add to the elements of the third row the products of the elements of the first, second and fourth rows by $-b_0$, $-b_1$, a_1 respectively. Add to the elements of the fourth row the products of those of the second by $-b_0$. We get

$$a_0^2\, R = \begin{vmatrix} a_0 & a_1 & a_2 & a_3 & a_4 & 0 \\ 0 & a_0 & a_1 & a_2 & a_3 & a_4 \\ 0 & 0 & (a_0b_2) & (a_1b_2) - a_3b_0 & -a_3b_1 - a_4b_0 & -a_4b_1 \\ 0 & 0 & (a_0b_1) & (a_0b_2) & -a_3b_0 & -a_4b_0 \\ 0 & 0 & b_0 & b_1 & b_2 & 0 \\ 0 & 0 & 0 & b_0 & b_1 & b_2 \end{vmatrix}.$$

Hence R equals the minor enclosed by dots.

EXERCISES †

1. For $m = 3$, $n = 2$, apply to Sylvester's determinant R exactly the same operations as used in the last case in § 8 and obtain

$$R = \begin{vmatrix} (a_0b_2) & (a_1b_2) - a_3b_0 & -a_3b_1 \\ (a_0b_1) & (a_0b_2) & -a_3b_0 \\ b_0 & b_1 & b_2 \end{vmatrix}.$$

2. Hence show that the discriminant of $a_0x^3 + a_1x^2 + a_2x + a_3 = 0$ is

$$-\begin{vmatrix} 2\,a_0a_2 & a_1a_2 + 3\,a_0a_3 & 2\,a_1a_3 \\ a_1 & 2\,a_2 & 3\,a_3 \\ 3\,a_0 & 2\,a_1 & a_2 \end{vmatrix}$$

$$= 18\,a_0a_1a_2a_3 - 4\,a_0a_2^3 - 4\,a_1^3a_3 + a_1^2a_2^2 - 27\,a_0^2a_3^2.$$

3. For $m = n = 4$, reduce Sylvester's R (as in the first case in § 8) to

$$\begin{vmatrix} (a_0b_1) & (a_0b_2) & (a_0b_3) & (a_0b_4) \\ (a_0b_2) & (a_0b_3) + (a_1b_2) & (a_0b_4) + (a_1b_3) & (a_1b_4) \\ (a_0b_3) & (a_0b_4) + (a_1b_3) & (a_1b_4) + (a_2b_3) & (a_2b_4) \\ (a_0b_4) & (a_1b_4) & (a_2b_4) & (a_3b_1) \end{vmatrix}.$$

4. For f and g of degree n, the ith function (8), when written as a determinant of the second order, is seen to equal

$$d_{i1}x^{n-1} + d_{i2}x^{n-2} + \cdots + d_{in},$$

where

$$d_{ij} = (a_0b_{i+j-1}) + (a_1b_{i+j-2}) + \cdots + (a_{i-1}b_j).$$

Then

$$R = (-1)^{\frac{n(n-1)}{2}} D, \quad D = \begin{vmatrix} d_{11} & \cdots & d_{1n} \\ \cdot & \cdot & \cdot \\ d_{n1} & \cdots & d_{nn} \end{vmatrix}.$$

This D is called the Bézout determinant of f and g. Show that $d_{ji} = d_{ij}$.

5. Hence verify for $m = n = 5$ that R can be derived from

$$\begin{vmatrix} (a_0b_1) & (a_0b_2) & (a_0b_3) & (a_0b_4) & (a_0b_5) \\ (a_0b_2) & (a_0b_3) & (a_0b_4) & (a_0b_5) & (a_1b_5) \\ (a_0b_3) & (a_0b_4) & (a_0b_5) & (a_1b_5) & (a_2b_5) \\ (a_0b_4) & (a_0b_5) & (a_1b_5) & (a_2b_5) & (a_3b_5) \\ (a_0b_5) & (a_1b_5) & (a_2b_5) & (a_3b_5) & (a_4b_5) \end{vmatrix}$$

by adding to its nine central elements the elements of

$$\begin{vmatrix} (a_1b_2) & (a_1b_3) & (a_1b_4) \\ (a_1b_3) & (a_1b_4) + (a_2b_3) & (a_2b_4) \\ (a_1b_4) & (a_2b_4) & (a_3b_4) \end{vmatrix}.$$

6. If $R(f, g) = 0$, we obtain a consistent set of equations by omitting one of Bézout's equations. Hence they determine x. If $m = n = 2$, find x. If $m = n = 3$, find x.

7. If $m \geqq n$, set $g_1(x) = x^{m-n}g(x)$. Then

$$R(f, g) = R(f, g_1) \div (-1)^{m(m-n)}a_m{}^{m-n}.$$

8. If $m = n$, $R(cf + dg, sf + tg) = \pm(ct - ds)^m R(f, g)$. [Find the new (a_ib_j).]

9. Express as a determinant of order m the resultant of $f(x) = 0$ and $x^m = 1$. [Multiply f by x and reduce by $x^m = 1$; repeat.]

9.† Without employing the results of §§ 3, 4, we may give a direct proof that the determinant (5) is the resultant of f and g, given by (1). While the method is general, we shall present it only in the case $m = 3$, $n = 2$. In the equation

$$(9) \qquad \begin{vmatrix} a_0 & a_1 & a_2 & a_3 - z & 0 \\ 0 & a_0 & a_1 & a_2 & a_3 - z \\ b_0 & b_1 & b_2 & 0 & 0 \\ 0 & b_0 & b_1 & b_2 & 0 \\ 0 & 0 & b_0 & b_1 & b_2 \end{vmatrix} = 0,$$

take $z = f(\beta_i)$. Multiply the elements of the first four columns by β_i^4, β_i^3, β_i^2, β_i, respectively, and add the products to the last column. All of the elements of the new last column are zero. Hence $f(\beta_1)$ and $f(\beta_2)$ are the roots of (9). Since the equation is of the form

$$b_0^3z^2 + (\quad)z + F = 0,$$

where F is given by (4), we have

$$F = b_0^3 f(\beta_1) f(\beta_2).$$

Hence the Sylvester determinant F is the resultant $R(f, g)$.

Moreover, the equation in z is the eliminant of

$$g(x) = 0, \quad z = f(x),$$

and hence gives explicitly the equation obtained from $g(x) = 0$ by applying the transformation $z = f(x)$ of Tschirnhausen (Ch. VII, § 13).

10.† Theorem. *Necessary and sufficient conditions that $f(x)$ and $g(x)$ shall have a common divisor of degree d, but none of higher degree, are $R = 0$, $R_1 = 0, \ldots, R_{d-1} = 0$, $R_d \neq 0$, where R is the determinant (5), and R_k is the determinant derived from R by deleting the last k rows of a's, the last k rows of b's, and the last $2k$ columns.*

For example, if $m = n = 4$,

$$(10) \qquad R_1 = \begin{vmatrix} a_0 & a_1 & a_2 & a_3 & a_4 & 0 \\ 0 & a_0 & a_1 & a_2 & a_3 & a_4 \\ 0 & 0 & a_0 & a_1 & a_2 & a_3 \\ b_0 & b_1 & b_2 & b_3 & b_4 & 0 \\ 0 & b_0 & b_1 & b_2 & b_3 & b_4 \\ 0 & 0 & b_0 & b_1 & b_2 & b_3 \end{vmatrix}.$$

To prove the theorem for the case $d = 1$, set

$$f_1 = p_1 x^{m-2} + \cdots + p_{m-1}, \quad g_1 = q_1 x^{n-2} + \cdots + q_{n-1}.$$

The conditions for an identity of the form

$$(11) \qquad f g_1 - g f_1 \equiv cx + c'$$

are

$$\begin{array}{llll} a_0 q_1 & - b_0 p_1 & & = 0, \\ a_1 q_1 + a_0 q_2 & - b_1 p_1 - b_0 p_2 & & = 0, \\ \ldots & \ldots & & \\ & a_m q_{n-2} + a_{m-1} q_{n-1} & - b_n p_{m-2} - b_{n-1} p_{m-1} & = c, \\ & a_m q_{n-1} & - b_n p_{m-1} & = c'. \end{array}$$

Omitting the last equation, we have $m + n - 2$ linear equations for the same number of unknowns q_i, $-p_i$. The determinant of the coefficients equals R_1 with the rows and columns interchanged. Hence if $R_1 \neq 0$ we may choose $c = R_1$ and find values not all zero of the unknowns satisfying all of the above equations except the last, and then choose c' so that the last holds. Let $R = 0$. Then f and g have a common linear factor, but no common factor of degree > 1 since the right member of (11) is of degree unity.

But if $R = R_1 = 0$, we may take $c = 0$ and find values not all zero of q_i, p_i satisfying all but the last of the above equations. The resulting value of c' is zero by (11), with $c = 0$, since f and g have a common factor $x - r$. Then

$$\frac{f}{x - r} g_1 - \frac{g}{x - r} f_1 \equiv 0.$$

Since not all of the $m - 1$ linear factors of the first fraction are factors of f_1 (of degree $m - 2$), at least one is a factor of the second fraction. Hence if $R = R_1 = 0$, f and g have a common factor of degree > 1.

To prove the theorem for $d = 2$, we employ functions f_2 and g_2 of degrees $m - 3$ and $n - 3$, respectively. Of the conditions for the identity

$$fg_2 - gf_2 \equiv cx^2 + c'x + c'', \tag{12}$$

we omit the two in which c' and c'' occur and see that the determinant of the coefficients of the remaining equations is R_2. Then if

$$R = R_1 = 0, \qquad R_2 \neq 0,$$

we may take $c = R_2$ and satisfy all of the conditions for (12). Thus f and g have no common factor of degree > 2.

EXERCISES †

1. By performing on (10) exactly the same operations as used in § 8 to reduce a determinant of order 6 to one of order 3, show that

$$R_1 = \begin{vmatrix} (a_0b_3) & (a_0b_4) + (a_1b_3) & (a_1b_4) + (a_2b_3) \\ (a_0b_2) & (a_0b_3) + (a_1b_2) & (a_0b_4) + (a_1b_3) \\ (a_0b_1) & (a_0b_2) & (a_0b_3) \end{vmatrix}.$$

Note that if $a_4 = b_4 = 0$, the present work reduces to the former.

2. In the notation of Ex. 4, p. 162, the preceding R_1 with its first and third rows interchanged becomes D_1:

$$D_1 = \begin{vmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{vmatrix}, \qquad R_1 = -D_1.$$

3. For $m = n$,

$$R_k = (-1)^{(n-k)(n-k-1)/2} D_k, \quad D_k = \begin{vmatrix} d_{11} & \cdots & d_{1\,n-k} \\ \cdot & \cdots & \cdot \\ d_{n-k\,1} & \cdots & d_{n-k\,n-k} \end{vmatrix}.$$

4. Hence, if $m = n$, f and g have a common divisor of degree d, but none of degree $> d$, if and only if $D = 0$, $D_1 = 0, \ldots, D_{d-1} = 0$, $D_d \neq 0$.

5. Give a direct proof of Ex. 4 by multiplying the ith function in Ex. 4, p. 162, by a variable y_i and summing for $i = 1, \ldots, t$. Thus

$$g \cdot \{a_0 y_1 + (a_0 x + a_1) y_2 + \cdots + (a_0 x^{t-1} + \cdots) y_t\} - f \cdot \{b_0 y_1 + (b_0 x + b_1) y_2 + \cdots\}$$
$$\equiv \delta_1 x^{n-1} + \delta_2 x^{n-2} + \cdots + \delta_n,$$

where $\qquad \delta_1 = d_{11} y_1 + \cdots + d_{t\,1} y_t, \ldots, \delta_n = d_{1\,n} y_1 + \cdots + d_{t\,n} y_t.$

The determinant of the coefficients of $y_1, \ldots, y_t$ in $\delta_1, \ldots, \delta_t$ is D_{n-t}. If $D = 0$, take $t = n$; then we can choose $y_1, \ldots, y_n$ not all zero so that $\delta_1 = 0, \ldots, \delta_n = 0$. Then $gf_1 - fg_1 \equiv 0$ for functions f_1 and g_1 of degree $n - 1$, so that f and g have a linear divisor. If also $D_1 = 0$, take $t = n - 1$; then we can make $\delta_1 = 0, \ldots, \delta_{n-1} = 0$. Hence $gf_2 - fg_2 \equiv \delta_n$ for functions f_2 and g_2 of degree $n - 2$. Since f and g have a common divisor, the constant δ_n is zero, and hence they have a common divisor of degree $\geqq 2$. But if $D_1 \neq 0$, we can make

$$gf_2 - fg_2 \equiv \delta_{n-1} x + \delta_n, \quad \delta_{n-1} \neq 0,$$

so that the only common divisor is linear.

MISCELLANEOUS EXERCISES

1. Find a necessary and sufficient condition that the roots α, β, γ of $x^3 + px^2 + qx + r = 0$ shall be in geometrical progression.

2. For the same equation find $\Sigma\alpha^3\beta^3$. [Replace x by $1/x$.]

3. Find the equation with the roots $\alpha^2 + \beta^2$, $\alpha^2 + \gamma^2$, $\beta^2 + \gamma^2$.

4. Find the equation with the roots $\alpha^2 + \beta^2 - \gamma^2$, $\alpha^2 + \gamma^2 - \beta^2$, etc.

5. Find the equation with the roots $\alpha^2 + \alpha\beta + \beta^2$, etc.

6. Solve the equation in Ex. 1 by forming and solving the quadratic equation with the roots $(\alpha + \omega\beta + \omega^2\gamma)^3$ and $(\alpha + \omega^2\beta + \omega\gamma)^3$, where $\omega^2 + \omega + 1 = 0$. (Lagrange.)

7. Solve $x^3 - 28\,x + 48 = 0$, given that two roots differ by 2.

8. Find a necessary and sufficient condition that

$$f(x) \equiv x^4 + px^3 + qx^2 + rx + s = 0$$

shall have one root the negative of another root. When this condition is satisfied, what are the quadratic factors of $f(x)$?

9. Solve $f(x) \equiv x^4 - 6\,x^3 + 13\,x^2 - 14\,x + 6 = 0$, given that two roots α and β are such that $2\,\alpha + \beta = 5$. Hint: $f(x)$ and $f(5 - 2\,x)$ have a common factor.

10. Diminish the roots of $x^4 + qx^2 + rx + s = 0$ $(s \neq 0)$ by such a number that the roots of the transformed equation shall be of the form a, m/a, b, m/b, and show how the latter equation may be solved.

11. Solve $x^4 - 2\,x^2 - 16\,x + 1 = 0$ by the method of Ex. 10.

12. By use of the equation whose roots are the squares of the roots of $x^5 + x^3 - x^2 + 2\,x - 3 = 0$ and Descartes' rule, show that the latter equation has four imaginary roots.

13. Similarly, $x^3 + x^2 + 8\,x + 6 = 0$ has imaginary roots.

14. If all of the roots of $x^n + ax^{n-1} + bx^{n-2} + \cdots = 0$ are real,

$$a^2 - 2\,b > 0, \quad b^2 - 2\,ac + 2\,d > 0, \quad c^2 - 2\,bd + 2\,ae - 2\,f > 0, \; \ldots .$$

Hint: Form the equation in $y = x^2$.

15. Solve $x^3 + px + q = 0$ by eliminating x between it and $x^2 + vx + w = y$ by the greatest common divisor process, and choosing v and w so that in the resulting cubic equation for y the coefficients of y and y^2 are zero. The next to the last step of the elimination gives x as a rational function of y. (Tschirnhausen, *Acta Erudit.*, Lipsiae, II, 1683, p. 204.)

16. Find the preceding y-cubic as follows. Multiply $x^2 + vx + w = y$ by x and replace x^3 by $-px - q$; then multiply the resulting quadratic equation in x by x and replace x^3 by its value. The determinant of the coefficients of x^2, x, 1 must vanish.

17. Eliminate y between $y^3 = v$, $x = ry + sy^2$, and get

$$x^3 - 3\,rsvx - (r^3v + s^3v^2) = 0.$$

Take $s = 1$ and choose r and v so that this equation shall be identical with $x^3 + px + q = 0$, and hence solve the latter. (Euler, 1764.)

18. Eliminate y between $y^3 = v$, $x = f + ey + y^2$ and get

$$\begin{vmatrix} 1 & e & f - x \\ e & f - x & v \\ f - x & v & ev \end{vmatrix} = 0.$$

This cubic equation in x may be identified with the general cubic equation by choice of e, f, v. Hence solve the latter.

19. Determine r, s and v so that the resultant of

$$y^3 = v, \qquad y = \frac{x + r}{y + s}$$

shall be identical with $x^3 + px + q = 0$. (Bézout, 1762.)

20. Show that the reduction of a cubic equation in x to the form $y^3 = v$ by the substitution

$$x = \frac{r + sy}{1 + y}$$

is not essentially different from the method of Ex. 18. [Multiply the numerator and denominator of x by $1 - y + y^2$.]

21. If the discriminant of a cubic equation is positive, the number of positive roots equals the number of variations of signs of the coefficients.

22. Descartes' rule gives the exact number of positive roots only when all the coefficients are of like sign or when

$$f(x) = x^n + p_1x^{n-1} + \cdots + p_{n-s}x^s - p_{n-s+1}x^{s-1} - \cdots - p_n = 0,$$

each p_i being $\geqq 0$. Without using that rule, show that the latter equation has one and only one positive root r. Hints: There is a positive root r by Ch. I, § 12 ($a = 0$, $b = \infty$). Call $P(x)$ the quotient of the sum of the positive terms by x^s, and call $-N(x)$ that of the negative terms. Then $N(x)$ is a sum of powers of $1/x$ with positive coefficients.

If	$x > r$,	$P(x) > P(r)$,	$N(x) < N(r)$,	$f(x) > 0$;	
If	$x < r$,	$P(x) < P(r)$,	$N(x) > N(r)$,	$f(x) < 0$.	(Lagrange.)

23. If $f(x) = f_1(x) + \cdots + f_k(x)$, where each $f_j(x)$ is like the f in Ex. 22, and if R is the greatest of the single positive roots of $f_1 = 0, \ldots, f_k = 0$, then R is an upper limit to the positive roots of $f = 0$.

24. Any cubic or quartic equation in x can be transformed into a reciprocal equation by a substitution $x = ry + s$.

25. Admitting that an equation $f(x) \equiv x^n + \cdots = 0$ with real coefficients has n roots, show algebraically that there is a real root between a and b if $f(a)$ and $f(b)$ have opposite signs. Note that a pair of conjugate imaginary roots $c \pm di$ are the roots of $(x - c)^2 + d^2 = 0$ and that this quadratic function is

positive if x is real. Hence if $x_1, \ldots, x_r$ are the real roots and $\phi(x) \equiv (x - x_1) \ldots (x - x_r)$, then $\phi(a)$ and $\phi(b)$ have opposite signs. Thus $a - x_i$ and $b - x_i$ have opposite signs for at least one real root x_i. (Lagrange.)

26. If s_j is the sum of the jth powers of the roots of an equation of degree n and if m is any integer, the equation is

$$\begin{vmatrix} x^n & x^{n-1} & \ldots & x & 1 \\ s_{m+n} & s_{m+n-1} & \ldots & s_{m+1} & s_m \\ s_{m+n+1} & s_{m+n} & \ldots & s_{m+2} & s_{m+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ s_{m+2n-1} & s_{m+2n-2} & \ldots & s_{m+n} & s_{m+n-1} \end{vmatrix} = 0.$$

Hint: Use the second set of Newton's identities. (Jacobi.)

27. If $a < b < c \ldots < l$, and $\alpha, \beta, \ldots, \lambda$ are positive,

$$\frac{\alpha}{x-a} + \frac{\beta}{x-b} + \frac{\gamma}{x-c} + \cdots + \frac{\lambda}{x-l} + t = 0$$

has a real root between a and b, one between b and c, . . . , one between k and l, and if t is negative one greater than l, but if t is positive one less than a.

28. Verify that the equation in Ex. 27 has no imaginary root by substituting $r + si$ and $r - si$ in turn for x, and subtracting the results.

29. In the problem of three astronomical bodies occurs the equation

$$r^5 + (3 - \mu)r^4 + (3 - 2\mu)r^3 - \mu r^2 - 2\mu r - \mu = 0,$$

where $0 < \mu < 1$. Why is there a single positive real root? As μ approaches zero, two complex roots and the real root approach zero.

30. Discuss the equation obtained from the preceding by changing the signs of the coefficients of r^4 and r.

31. By Newton's identities,

$$s_3 = -\begin{vmatrix} 1 & 0 & p_1 \\ p_1 & 1 & 2p_2 \\ p_2 & p_1 & 3p_3 \end{vmatrix} = -p_1^3 + 3p_1p_2 - 3p_3,$$

$$s_k = -\begin{vmatrix} 1 & 0 & 0 & \ldots & 0 & p_1 \\ p_1 & 1 & 0 & \ldots & 0 & 2p_2 \\ p_2 & p_1 & 1 & \ldots & 0 & 3p_3 \\ p_3 & p_2 & p_1 & \ldots & 0 & 4p_4 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ p_{k-1} & p_{k-2} & p_{k-3} & \ldots & p_1 & kp_k \end{vmatrix},$$

where all but the last term in the main diagonal is 1, and all terms above the diagonal are zero except those in the last column. If $k > n$, we must set $p_j = 0 (j > n)$.

32. By Newton's identities,

$$3!\,p_3 = -\begin{vmatrix} 1 & 0 & s_1 \\ s_1 & 2 & s_2 \\ s_2 & s_1 & s_3 \end{vmatrix}, \quad k!\,p_k = -\begin{vmatrix} 1 & 0 & 0 & \dots & 0 & s_1 \\ s_1 & 2 & 0 & \dots & 0 & s_2 \\ s_2 & s_1 & 3 & \dots & 0 & s_3 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ s_{k-1} & s_{k-2} & s_{k-3} & \dots & k & s_k \end{vmatrix},$$

if $k \leqq n$. But if $k \geqq n$,

$$\begin{vmatrix} s_k & s_{k-1} & s_{k-2} & \dots & s_{k-n} \\ s_{k+1} & s_k & s_{k-1} & \dots & s_{k-n+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ s_{k+n} & s_{k+n-1} & s_{k+n-2} & \dots & s_k \end{vmatrix} = 0.$$

33. Let $s_i = \alpha_1{}^i + \cdots + \alpha_n{}^i$. Let $\alpha_1^2, \dots, \alpha_n^2$ be the roots of

$$y^n + P_1 y^{n-1} + \cdots + P_n = 0.$$

Set $y = \alpha_j^2$ and multiply the result by α_j^{k-2n}, where $k \geqq 2n$. Sum for $j = 1, \dots, n$. Thus

$$s_k + P_1 s_{k-2} + P_2 s_{k-4} + \cdots + P_n s_{k-2n} = 0.$$

Hence

$$\begin{vmatrix} s_k & s_{k-2} & s_{k-4} & \dots & s_{k-2n} \\ s_{k+1} & s_{k-1} & s_{k-3} & \dots & s_{k-2n+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ s_{k+n} & s_{k+n-2} & s_{k+n-4} & \dots & s_{k-n} \end{vmatrix} = 0.$$

34. Obtain a vanishing determinant similar to that in Ex. 33 but having the subscripts of the s's in each row decreased by 3.

35.
$$\begin{vmatrix} s_0 & s_1 & s_2 \\ s_1 & s_2 & s_3 \\ s_2 & s_3 & s_4 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & s_0 & s_1 & s_2 \\ s_0 & s_1 & s_2 & s_3 \\ s_1 & s_2 & s_3 & s_4 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & p_1 & p_2 & p_3 \\ 0 & s_0 & s_1 + p_1 s_0 & s_2 + p_1 s_1 + p_2 s_0 \\ s_0 & s_1 + p_1 s_0 & s_2 + p_1 s_1 + p_2 s_0 & s_3 + p_1 s_2 + p_2 s_1 + p_3 s_0 \\ s_1 & s_2 + p_1 s_1 & s_3 + p_1 s_2 + p_2 s_1 & s_4 + p_1 s_3 + p_2 s_2 + p_3 s_1 \end{vmatrix}$$

$$= -\begin{vmatrix} 1 & p_1 & p_2 & p_3 \\ 0 & n & (n-1)p_1 & (n-2)p_2 \\ n & (n-1)p_1 & (n-2)p_2 & (n-3)p_3 \\ p_1 & 2\,p_2 & 3\,p_3 & 4\,p_4 \end{vmatrix}.$$

36. If $n = 3$, the last determinant may be obtained from the Sylvester resultant R of $x^3 + p_1x^2 + p_2x + p_3$ and its derivative by multiplying the elements of the first row of R by -3 and adding the products to the elements of the third row.

37. Express the determinant of order 4 in the s_i (analogous to the first one in Ex. 35) as a determinant of order 6 in the p's. For $n = 4$, identify the latter with the resultant of $x^4 + p_1x^3 + p_2x^2 + p_3x + p_4$ and its derivative.

38. Let s_k be the sum of the kth powers of the roots $x_1, \ldots, x_n$ of a given equation. The coefficients of the equation having as its roots the $\frac{1}{2}n(n-1)$ squares of the differences of the x's can be found from $S_1, S_2, \ldots$, where S_p is the sum of the pth powers of the roots of the latter equation. Expand by the binomial theorem

$$(x - x_1)^{2p} + (x - x_2)^{2p} + \cdots + (x - x_n)^{2p},$$

set $x = x_1, \ldots, x = x_n$ in turn, add and divide by 2. Thus

$$S_p = ns_{2p} - 2ps_{2p-1}s_1 + \frac{(2p)(2p-1)}{1 \cdot 2} s_{2p-2}s_2 - \cdots \pm \tfrac{1}{2} \frac{2p(2p-1)\ldots(p+1)}{1 \cdot 2 \ldots p} s_p^2.$$

(Lagrange.)

39. In particular, $S_1 = ns_2 - s_1^2$, $S_2 = ns_4 - 4s_1s_3 + 3s_2^2$, $S_3 = ns_6 - 6s_1s_5 + 15s_2s_4 - 10s_3^2$. Hence give the equation whose roots are the squares of the differences of the roots of a given cubic equation. Deduce the discriminant of the latter.

40. The equation whose roots are the $n(n-1)$ differences $x_j - x_k$ of the roots of $f(x) = 0$ may be obtained by eliminating x between the latter and $f(x + y) = 0$ and deleting the factor y^n (arising from $y = x_j - x_j = 0$) from the eliminant. The equation free of this factor may be obtained by eliminating x between $f(x) = 0$ and

$$\{f(x+y) - f(x)\}/y = f'(x) + f''(x)\frac{y}{1 \cdot 2} + \cdots + f^n(x)\frac{y^{n-1}}{1 \cdot 2 \ldots n} = 0.$$

This eliminant involves only even powers of y, so that if we set $y^2 = z$ we obtain an equation in z having as its roots the squares of the differences of the roots of $f(x) = 0$. (Lagrange.)

41. Compute by Ex. 40 the z-equation when $f(x) = x^3 + px + q$.

42. Except for $b = 0$, the equation

$$\begin{vmatrix} a - x & b \\ b & f - x \end{vmatrix} = 0,$$

has a real root exceeding a and f, and one less than a and f. [Substitute a and f for x in turn].

43. Let the equation in Ex. 42 have distinct real roots α, β, where $\alpha > \beta$. Then there are three real roots of*

$$D(x) \equiv \begin{vmatrix} a-x & b & c \\ b & f-x & g \\ c & g & h-x \end{vmatrix} = 0.$$

Hint: The results of substituting α and β for x in $D(x)$ are

$$\left[c\sqrt{\alpha - f} + g\sqrt{\alpha - a}\right]^2, \qquad -\left[c\sqrt{f-\beta} - g\sqrt{a-\beta}\right]^2,$$

where the product of the radicals in each is $+b$. Hence if neither α nor β is a root, there is a root $> \alpha$, one $< \beta$, and one between α and β. If α is a root, there is a root $< \beta$ and hence three real roots.

44. If $\alpha = \beta$ in Ex. 43, then $a = f$ is a root of $D(x) = 0$ and there are two further real roots.

45.
$$\begin{vmatrix} aa' + bb' + cc' & ea' + fb' + gc' \\ ae' + bf' + cg' & ee' + ff' + gg' \end{vmatrix} = af\begin{vmatrix} a' & b' \\ e' & f' \end{vmatrix} + ag\begin{vmatrix} a' & c' \\ e' & g' \end{vmatrix}$$
$$+ be\begin{vmatrix} b' & a' \\ f' & e' \end{vmatrix} + bg\begin{vmatrix} b' & c' \\ f' & g' \end{vmatrix} + ce\begin{vmatrix} c' & a' \\ g' & e' \end{vmatrix} + cf\begin{vmatrix} c' & b' \\ g' & f' \end{vmatrix}.$$

Combine the first and third, second and fifth, fourth and sixth:

$$\begin{vmatrix} a & b \\ e & f \end{vmatrix} \cdot \begin{vmatrix} a' & b' \\ e' & f' \end{vmatrix} + \begin{vmatrix} a & c \\ e & g \end{vmatrix} \cdot \begin{vmatrix} a' & c' \\ e' & g' \end{vmatrix} + \begin{vmatrix} b & c \\ f & g \end{vmatrix} \cdot \begin{vmatrix} b' & c' \\ f' & g' \end{vmatrix}.$$

46. Hence, in particular,

$$\begin{vmatrix} a^2 + b^2 + c^2 & ae + bf + cg \\ ae + bf + cg & e^2 + f^2 + g^2 \end{vmatrix} = \begin{vmatrix} a & b \\ e & f \end{vmatrix}^2 + \begin{vmatrix} a & c \\ e & g \end{vmatrix}^2 + \begin{vmatrix} b & c \\ f & g \end{vmatrix}^2.$$

47. Hence if a, b, c and e, f, g are the direction cosines of two lines in space, and if θ is the angle between them, so that $\cos\theta = ae + bf + cg$, then $\sin^2\theta$ equals the above sum of three squares.

48. For the determinant in Ex. 43,

$$D(x)\cdot D(-x) = \begin{vmatrix} a^2+b^2+c^2-x^2 & ab+bf+cg & ac+bg+ch \\ ab+bf+cg & b^2+f^2+g^2-x^2 & bc+fg+gh \\ ac+bg+ch & bc+fg+gh & c^2+g^2+h^2-x^2 \end{vmatrix}$$
$$= -x^6 + x^4(a^2+f^2+h^2+2b^2+2c^2+2g^2) - x^2(D_1+D_2+D_3) + D^2(0),$$

* This theorem is important in many branches of pure and applied mathematics. Besides this proof and that in Ex. 48, other more advanced proofs, including that by Borchardt, are given in Salmon's *Modern Higher Algebra*, pp. 48–56.

where D_3 is the first determinant in Ex. 46 for $e = b$ and D_1 and D_2 are analogous minors of elements in the main diagonal of the present determinant of order 3 with $x = 0$. Hence the coefficient of $-x^2$ is a sum of squares (Ex. 49). Since the function of degree 6 is not zero for a negative value of x^2, $D(x) = 0$ has no purely imaginary root. If it had an imaginary root $r + si$, then $D(x + r) = 0$ would have a purely imaginary root si. But $D(x + r)$ is of the form in Ex. 43 with a, f, h replaced by $a - r, f - r, h - r$. Hence $D(x) = 0$ has only real roots. The method is applicable to such determinants of order n. (Sylvester.)

49. In Ex. 48, $D_1 + D_2 + D_3$ equals

$$(af - b^2)^2 + (ah - c^2)^2 + (fh - g^2)^2 + 2\,(ag - bc)^2 + 2\,(cf - bg)^2 + 2\,(bh - cg)^2.$$

50. Without using its solution by radicals, prove that

$$x^4 + bx^3 + cx^2 + dx + e$$

has a factor $x^2 - sx + p$, where s is a root of a sextic equation, and that p is a rational function of s and the coefficients.

Hints: There are six functions like $s = x_1 + x_2$; next,

$$c = \Sigma x_1x_2 = s(x_3 + x_4) + p + x_3x_4,$$
$$-d = \Sigma x_1x_2x_3 = sx_3x_4 + (x_3 + x_4)p.$$

Replace $x_3 + x_4$ by $-b - s$ and solve for p the resulting linear equations in x_3x_4 and p. The case $b + 2\,s = 0$ may be avoided by starting with another pair of roots.

51. Prove Ex. 50 by dividing the quartic by the quadratic function and requiring that the linear remainder shall be zero identically.

52. Prove Ex. 50 by use of (3) and (8) in Ch. IV.

53. $x^6 + bx^5 + cx^4 + dx^3 + ex^2 + fx + g$ has a factor $x^2 - sx + p$, where s is a root of an equation of degree 15, and p is a rational function of s and the coefficients. Hints: Write

$$\sigma_1 = x_3 + x_4 + x_5 + x_6, \quad \sigma_2 = x_3x_4 + \cdots, \quad \sigma_3 = x_3x_4x_5 + \cdots, \quad \sigma_4 = x_3x_4x_5x_6$$

for the elementary symmetric functions of $x_3, \ldots, x_6$, and show that

$$-b = s + \sigma_1, \qquad c = p + s\sigma_1 + \sigma_2, \qquad -d = p\sigma_1 + s\sigma_2 + \sigma_3,$$
$$e = p\sigma_2 + s\sigma_3 + \sigma_4, \qquad -f = s\sigma_4 + p\sigma_3, \quad g = p\sigma_4.$$

The first four relations determine the σ's. Then the last two give a cubic and a quadratic equation in p, by means of which we may express p as a rational function of s and then obtain an equation in s alone. Why must this be of degree 15?

54. If Ex. 53 were solved as in Ex. 51 (if the quotient of $x^6 + \cdots$ by $x^2 + \cdots$ be denoted by $x^4 - \sigma_1x^3 + \sigma_2x^2 - \sigma_3x + \sigma_4$, we obtain the above six relations), why could we conclude that any equation of degree six with real coefficients has two complex roots (independently of the fundamental theorem of algebra)?

55. $$\begin{vmatrix} 3 & \alpha_1+\alpha_2+\alpha_3 \\ \alpha_1+\alpha_2+\alpha_3 & \alpha_1^2+\alpha_2^2+\alpha_3^2 \end{vmatrix} = \sum_3 (\alpha_1-\alpha_2)^2. \quad \text{[See Ex. 46.]}$$

56. The determinant in Ex. 55 equals

$$\sum_6 \begin{vmatrix} 1 & \alpha_2 \\ \alpha_1 & \alpha_2^2 \end{vmatrix} = \sum_3 \left\{ \begin{vmatrix} 1 & \alpha_2 \\ \alpha_1 & \alpha_2^2 \end{vmatrix} + \begin{vmatrix} 1 & \alpha_1 \\ \alpha_2 & \alpha_1^2 \end{vmatrix} \right\} = \sum_3 \begin{vmatrix} 2 & \alpha_1+\alpha_2 \\ \alpha_1+\alpha_2 & \alpha_1^2+\alpha_2^2 \end{vmatrix} = \sum_3 \begin{vmatrix} 1 & \alpha_1 \\ 1 & \alpha_2 \end{vmatrix}^2.$$

57. For n roots, $n \geqq 3$,

$$D \equiv \begin{vmatrix} n & s_1 & s_2 \\ s_1 & s_2 & s_3 \\ s_2 & s_3 & s_4 \end{vmatrix} = \sum \begin{vmatrix} 1 & \alpha_j & \alpha_k^2 \\ \alpha_i & \alpha_j^2 & \alpha_k^3 \\ \alpha_i^2 & \alpha_j^3 & \alpha_k^4 \end{vmatrix} \quad \begin{pmatrix} i, j, k = 1, \ldots, n \\ i, j, k \text{ distinct.} \end{pmatrix}.$$

Add the six determinants given by the permutations of fixed i, j, k. Then

$$D = \sum_{i<j<k} \begin{vmatrix} 1+1+1 & \alpha_i+\alpha_j+\alpha_k & \alpha_i^2+\alpha_j^2+\alpha_k^2 \\ \alpha_i+\alpha_j+\alpha_k & \alpha_i^2+\alpha_j^2+\alpha_k^2 & \alpha_i^3+\alpha_j^3+\alpha_k^3 \\ \alpha_i^2+\alpha_j^2+\alpha_k^2 & \alpha_i^3+\alpha_j^3+\alpha_k^3 & \alpha_i^4+\alpha_j^4+\alpha_k^4 \end{vmatrix}$$

$$= \sum_{i<j<k} \begin{vmatrix} 1 & 1 & 1 \\ \alpha_i & \alpha_j & \alpha_k \\ \alpha_i^2 & \alpha_j^2 & \alpha_k^2 \end{vmatrix} \cdot \begin{vmatrix} 1 & \alpha_i & \alpha_i^2 \\ 1 & \alpha_j & \alpha_j^2 \\ 1 & \alpha_k & \alpha_k^2 \end{vmatrix} = \sum (\alpha_i-\alpha_j)^2(\alpha_i-\alpha_k)^2(\alpha_j-\alpha_k)^2.$$

58. Comparing the theorems in Exs. 55 and 57 and their extensions with Ex. 12, p. 102, we see the nature of a proof of Borchardt's Theorem: An equation of degree n with real coefficients and distinct roots has as many pairs of imaginary roots as there are changes in signs in the series

$$s_0 = n, \quad \begin{vmatrix} s_0 & s_1 \\ s_1 & s_2 \end{vmatrix}, \quad \begin{vmatrix} s_0 & s_1 & s_2 \\ s_1 & s_2 & s_3 \\ s_2 & s_3 & s_4 \end{vmatrix}, \ldots, \quad \begin{vmatrix} s_0 & s_1 & \ldots & s_{n-1} \\ s_1 & s_2 & \ldots & s_n \\ \cdot & \cdot & \cdot & \cdot \\ s_{n-1} & s_n & \ldots & s_{2n-2} \end{vmatrix}.$$

If two consecutive terms are zero, the theorem may fail, as $x^4 + 1 = 0$ shows. But it holds if an isolated zero occurs and is suppressed.

59. Denote the last series by $D_1 = s_0$, D_2, D_3, . . . , D_n. There are exactly r distinct roots of the given equation of degree n if and only if D_r is the last non-vanishing determinant of this series. For, as in Exs. 55–57, D_k is the sum of the various products of the squares of the differences of k of the roots $\alpha_1, \ldots, \alpha_n$. If $k > r$, each product involves two equal α's and hence $D_k = 0$. If $k = r$, the only term not zero is that involving the r distinct α's, so that $D_r \neq 0$. (L. Baur, *Math. Annalen*, vols. 50, 52.)

60. The n roots are all real and distinct if and only if $D_2, \ldots, D_n$ are all positive. (Weber, *Algebra*, 2d ed., I, p. 322.)

61. If each c_i is real and if the numbers

$$c_0, \quad \begin{vmatrix} c_0 & c_1 \\ c_1 & c_2 \end{vmatrix}, \ldots, \quad \begin{vmatrix} c_0 & c_1 & \ldots & c_n \\ c_1 & c_2 & \ldots & c_{n+1} \\ \cdot & \cdot & \cdot & \cdot \\ c_n & c_{n+1} & \ldots & c_{2n+2} \end{vmatrix}$$

are positive, all of the roots of

$$c_0 + c_1x + c_2x^2 + \cdots + c_{2n}x^{2n} = 0$$

are imaginary, and all but one of the roots of

$$c_0 + c_1x + c_2x^2 + \cdots + c_{2n+1}x^{2n+1} = 0$$

are imaginary. (Van Vleck, *Annals of Math.*, 4 (1903), p. 191.)

62. The results in Ex. 61 follow if the c_{2i} and $\begin{vmatrix} c_{2i} & c_{2i+1} \\ c_{2i+1} & c_{2i+2} \end{vmatrix}$ are all positive. (Kellogg, *Annals of Math.*, 9 (1907), p. 97.)

63. If the terms with negative coefficients in an equation of degree n are $-\alpha x^{n-a}$, $-\beta x^{n-b}$, $-\gamma x^{n-c}$, ..., no positive root exceeds the sum of the two largest of the numbers

$$\sqrt[a]{\alpha}, \quad \sqrt[b]{\beta}, \quad \sqrt[c]{\gamma}, \ldots. \qquad \text{(Lagrange.)}$$

64. In Ex. 63, no positive root exceeds the greatest of the numbers

$$\sqrt[a]{k\alpha}, \quad \sqrt[b]{k\beta}, \ldots,$$

where k is the number of the negative coefficients $-\alpha$, ... (Cauchy.)

65.* Define V_a as in Ch. IX, § 8, and let $f(a) \neq 0$, $f(b) \neq 0$. If $f(x) = 0$ has imaginary roots, $V_a - V_b$ cannot give the exact number of real roots in every interval $[a, b]$; but, if $f(x) = 0$ has no imaginary roots, $V_a - V_b$ gives the exact number of real roots in every interval $[a, b]$. Hint: Use (14), Ch. IX.

66. Budan's Theorem gives the exact number of real roots of $f(x) = 0$ in $[a, b]$ if $f(a) \neq 0, f(b) \neq 0$, provided that, for $r = 0, 1, \ldots, n-2$, real roots of $f^{(r)}(x) = 0$ separate those of $f^{(r+1)}(x) = 0$ in that interval from each other and from a and b. The term "separate" here excludes the case of coincidence. Hint: At a root of $f^{(r+1)}(x) = 0$, the functions $f^{(r)}(x)$ and $f^{(r+2)}(x)$ must be of opposite sign.

67. Descartes' Rule gives the exact number of real roots only when Budan's Rule is exact for every positive interval $[a, b]$. Thus it is exact for an equation having only real roots.

68. We define as generalized Sturm's functions for an interval $[a, b]$ a sequence of polynomials $f(x), f_1(x), \ldots, f_r(x)$, with the following properties:

* The author is indebted to Professor D. R. Curtiss for Exs. 65–72.

(a) No two consecutive functions vanish simultaneously at any point of $[a, b]$;

(b) $f_r(x)$ does not vanish in $[a, b]$;

(c) When, for $1 \leqq i \leqq r - 1$, $f_i(x)$ vanishes for a value of x_1 in $[a, b]$, $f_{i-1}(x_1)$ and $f_{i+1}(x_1)$ have opposite signs;

(d) When $f(x)$ vanishes for a value x_1 in $[a, b]$, $f_1(x_1)$ has the same sign as $f'(x_1)$.

Prove that the number of real roots of $f(x) = 0$ in $[a, b]$ is equal to the difference between the numbers of variations of signs in such a sequence for $x = a$ and for $x = b$.

Prove the corresponding statement for an interval $[c, d]$ within $[a, b]$.

69. Prove that generalized Sturm's functions for any interval $[a, b]$, where a and b are both positive or both negative and $f(x) = 0$ has no multiple roots, may be obtained as follows: Take $f_1(x) = f'(x)$. Arrange $f(x)$ and $f_1(x)$ in ascending powers of x, and divide the former by the latter (using negative powers of x in the quotient, if necessary); let the last remainder of degree equal to that of $f(x)$ be designated by $r_2(x)$; then $f_2(x) = -r_2(x) \div x^2$. Define $f_i(x)$ similarly by division of $f_{i-2}(x)$ by $f_{i-1}(x)$, both being arranged according to ascending powers of x; the last remainder of degree equal to that of $f_{i-2}(x)$ is divided by $-x^2$ and the quotient taken as $f_i(x)$. Show that the sequence thus obtained is valid for $[-\infty, \infty]$, provided no one of the functions vanishes for $x = 0$.

70. Prove that generalized Sturm's functions for any interval $[a, b]$, where a and b are both positive or both negative and $f(x)=0$ has no multiple roots, may be obtained by the greatest common divisor process for $f(x) \equiv a_0x^n+a_1x^{n-1}+ \cdots +a_n$ and $f_1(x)$, with the signs of the remainders changed (as in Sturm's method), if we take

$$f_1(x) = \phi(x) \equiv a_1x^{n-1} + 2\,a_2x^{n-2} + \cdots + na_n \qquad (x < 0),$$

but $\quad f_1(x) = -\phi(x)$ if $x > 0$. $\quad$ Hint: $xf'(x) + \phi(x) = nf(x)$.

71. Prove the analogue of Ex. 69 when $f_1(x)$ is taken as in Ex. 70.

72. For the cubic $f(x) \equiv a_0x^3 + a_1x^2 + a_2x + a_3$ without multiple roots, discuss the validity of the sequences in Exs. 69–71 for any interval $[a, b]$, where $a < 0$, $b > 0$. Hint: If $a_3 \neq 0$, discuss whether variations of signs for x very near zero and negative = variations of signs for x very near zero and positive.

ANSWERS

Page 2.

1. 1.6, 4.4. 2. No real. 4. 1.2, -1.8, -3.4.

Page 7.

2. 2.1. 3. $(-0.845, 4.921)$, $(-3.155, 11.079)$; between -4 and -5.
4. 1.1, -1.3. 5. Between 0 and 1, 0 and -1, 2.5 and 3, -2.5 and -3.
9. $120(x^3 + x)$, $120x^2 - 42$.

Page 9.

1. 3. 2. 2, -2. 3. -1. 4. Double roots 1, 3.

Page 10.

3. Use Ex. 3, p. 9, abscissas -1, 3. 4. Use Ex. 2, p. 9.

Page 11.

1. One. 2. Three. 3. Three. 4. 1, 1, -2. 5. One.

Page 16.

7. 0.3, 1.5, -1.8. 8. 1.2. 9. 1.3, 1.7, -3.0. 15. 1, 2, 3, -6.

Page 23.

3. $\frac{1}{13}(19i - 9)$, $\dfrac{a^2 - b^2 + 2abi}{a^2 + b^2}$, $\frac{1}{5}(6 + \sqrt{5} - 3i + 2\sqrt{5}i)$.

4. Commutative and associative laws of addition and multiplication. Distributive law.

Page 24.

1. $\pm(3 + 4i)$. 2. $\pm(5 + 6i)$. 3. $\pm(3 - 2i)$.
4. $\pm[c + d + (c - d)i]$. 5. $\pm(c - di)$.

Page 26 (middle).

2. -3, -3ω, $-3\omega^2$; i, ωi, $\omega^2 i$.
3. $\cos A + i \sin A \quad (A = 40°, 160°, 280°)$.

Page 26 (bottom).

1. $-1, \cos A + i \sin A \quad (A = 36°, 108°, 252°, 324°)$.

Page 30.

2. $5, -1 \pm \sqrt{-3}$.
3. $1 \pm i, 1 \pm \sqrt{2}$.
4. $x^3 - 7x^2 + 19x - 13 = 0$.
5. $x^3 + (1+i)x^2 + 1 = 0$.
7. $\pm 1, 2 \pm \sqrt{3}$.
8, 9. $\sqrt{3}, 2 \pm i$.
10. $x^3 - \frac{3}{2}x^2 - \frac{5}{4}x + \frac{7}{8} = 0$.

Page 32.

2. $-5, \frac{1}{2}(5 \pm \sqrt{-3})$.
3. $-6, \pm\sqrt{-3}$.
4. $-2, 1 \pm i$.
5. $\frac{1}{4}, \frac{1}{7}(-2 \pm \sqrt{-3})$.

Page 34.

1. Three.
2. Two.
3. Two.

Page 35.

1. $-4, 2 \pm \sqrt{3}; 3, 3, -6$.
2. Page 37.
3. $1.3569, 1.6920, -3.0489$.

Page 37.

1. See 3, p. 35.
2. $-1.2017, 1.3300, -3.1284$.
3. $1.24698, -1.80194, -0.44504$.
4. $1.1642, -1.7728, -3.3914$.

Page 39.

2. $-1, -2, 2, 3$.
3. $1, -1, 4 \pm \sqrt{6}$.

Page 43.

2. $1 \pm \sqrt{2}, -1 \pm \sqrt{-2}$.
3. $4, -2, -1 \pm i$.
4. See Ex. 3, p. 39.

Page 53.

1. $z = -1 - 2i, \omega z, \omega^2 z$.

Page 56.

1. $x^4 - 8x^2 + 16 = 0$.
2. $1, 3$.
3. $4, 1 - \sqrt{-5}$.
5. $2 + \sqrt{3}, x^2 + 2x + 2 = 0$.
6. $1, 2$.
7. $-3, 1, 5$.
8. $4, \frac{3}{2}, -\frac{3}{2}$.
9. $1, 3, 5$.
10. $2, -6, 18$.
11. $5, 2, -1, -4$.
12. $1, 1, 1, 3$.
13. $p_3 = p_1 p_2$.
16. $y^3 - 12y - 12 = 0$.

Page 58.

3. 6.
4. 2.
5. 3.

Page 61.

1. $1, 3, 6.$ 2. $2, -1, -4, 5.$ 3. $-12, -35.$ 4. $2, 2, -3.$

Page 62.

1. $\frac{1}{3}, 1, 3, 9.$ 2. $1, \frac{1}{2}, \frac{1}{3}.$ 3. $-\frac{1}{6}.$ 4. $-\frac{1}{4}, -\frac{1}{4}, \frac{1}{2}.$

Page 65.

1. $q^2 - 2pr + 2s.$ 2. $p^2q - 2q^2 - pr + 4s.$
3. $p^4 - 4p^2q + 2q^2 + 4pr - 4s.$
4. $y^3 - (p^2 - 2q)y^2 + (q^2 - 2pr)y - r^2 = 0.$
5. $y^3 - qy^2 + pry - r^2 = 0.$ 6. $ry^3 + 2qy^2 + 4py + 8 = 0.$
7. $E_1^2 - 2E_2.$ 8. $E_1E_2 - 3E_3.$ 9. $E_1E_2.$
10. $E_1^3 - 3E_1E_2 + 3E_3.$ 11. $E_1^3 - 3E_1E_2.$
12. $E_1E_3 - 4E_4$ if $n > 3$, E_1E_3 if $n = 3$.

Page 71.

2. $s_a s_b s_c s_d - \Sigma s_a s_b s_{c+d} + 2\Sigma s_a s_{b+c+d} + \Sigma s_{a+b} s_{c+d} - 6 s_{a+b+c+d}$, if a, b, c, d are distinct; but if all are equal,

$$\tfrac{1}{24}(s_a^4 - 6s_a^2 s_{2a} + 8 s_a s_{3a} + 3 s_{2a}^2 - 6 s_{4a}).$$

3. See Exs. 1, 2, 12, 13, page 65.

Page 76.

3. $y^7 - 7qy^5 + 14q^2y^3 - 7q^3y = c.$ 4. $\epsilon^m - 2\epsilon^{5-m}$ $(m = 0, \ldots, 4).$

Page 77.

2. $\dfrac{p^4 - 3p^2q + 5pr + q^2}{r - pq}.$ 7. $2p^2 - 2q.$ 8. $-p^3 + 24r.$

9. $\dfrac{3p^2q^2 - 4p^3r - 4q^3 - 2pqr - 9r^2}{(r - pq)^2}.$

10. $27r^2 - 9pqr + 2q^3 = 0.$ 12. $y = q + r/x.$ 13. $x = \dfrac{1 - py}{2 + 2y}.$

Page 83.

1. $1, -\frac{1}{2}(1 \pm \sqrt{-3}), \frac{1}{2}(7 \pm \sqrt{45}).$ 2. $1, x^2 + \frac{1}{2}(1 \pm \sqrt{5})x + 1 = 0.$
3. $\pm 1, x^2 \pm x + 1 = 0.$ 5. $z^3 + z^2 - 2z - 1 = 0.$
6. $z^5 + z^4 - 4z^3 - 3z^2 + 3z + 1 = 0.$ 8. $2\cos 2\pi/7$, etc.

Page 88.

2. $g = 2, \quad r + r^8 + r^{12} + r^5$, etc., $\quad z^3 + z^2 - 4z + 1 = 0$.

Page 98.

1. One, between -2 and -3.
2. One, between 1 and 2.

Page 99.

1. Ex. 2, p. 37.
2. (0, 1), (-1.1, -1).
3. $x = -y$ in Ex. 1, p. 37.
4. (0, 1), (-2, -1).
5. (1, 2), (-7, -6).
6. (0, 1), (3, 4).

Page 103.

1. 2, -2.
4. 1, 1, two imaginary.

Page 105.

1. (-2, -1), (0, 1), (1, 2).
2. (-4, -3), (-2, -1), (1, 2).

Page 113.

1. Page 117.
2, 3. Exs. 1, 3, page 119.

Page 119.

1. -1.7728656.
2. $y = -x$ in Ex. 3, p. 35.
3. Single, -2.46954.
4, 5. Exs. 2, 3, p. 37.
6. Two negative and 2.121+, 2.123+.
7. 3.45592, 21.43067.
8. 2.24004099.

Page 121.

2. Darwin's quartic: $-12.4433 \pm 19.7596\, i$.
3. -0.59368, -2.04727, $1.32048 \pm 2.0039\, i$.
4. 0.35098, 12.75644, 32.0602, 34.8322.

Pages 123–24.

1. 4.0644364, $\quad -0.89196520$, $\quad$ 0.82752156.
2. § 1, $\quad -1.04727 \pm 1.13594\, i$.
5. -2.46955.

Page 128.

1. $x = 5, \quad y = 6$.
2. $x = 2, \quad y = 1$.
3. $x = a, \quad y = 0$.

Page 130.

1. $x = -8, \quad y = -7, \quad z = 26.$ 2. $x = 3, \quad y = -5, \quad z = 2.$

Page 134.

2. $x = 6, \quad y = 3, \quad z = 12.$ 3. $x = 5, \quad y = 4, \quad z = 3.$

4. $x = \dfrac{(k-b)(c-k)}{(a-b)(c-a)}.$ 5. $x = \dfrac{k(b-k)(c-k)(k+b+c)}{a(b-a)(c-a)(a+b+c)}.$

6. $x = \dfrac{a-1}{a+2}, \quad y = z = \dfrac{-3}{a+2}$, if $a \neq -2, 1.$

Page 137.

1. $-a_2b_1c_3d_4 + a_2b_1c_4d_3 + a_2b_3c_1d_4 - a_2b_3c_4d_1 - a_2b_4c_1d_3 + a_2b_4c_3d_1.$
2. $+, \quad +.$

Page 148.

2. Consistent: $y = -8/7 - 2x, \quad z = 5/7$ (common line).
3. Inconsistent, case (β).
4. Inconsistent (two parallel planes).
5. Consistent (single plane).

Page 149.

1. $x : y : z = -4 : 1 : 1.$ 2. $x : y : z = -10 : 8 : 7.$
3. Two unknowns arbitrary. 4. $x : y : z : w = 6 : 3 : 12 : 1.$
5. $z = -\dfrac{11}{3}x - \dfrac{19}{3}y, \quad w = -\dfrac{10}{3}x - \dfrac{17}{3}y.$
6. $y = -8/7 - 2x, \quad z = 5/7.$
7. Inconsistent (determinant 4th order $\neq 0$).

Page 167.

1. $p^3r = q^3.$ 2. $3r^2 - 3pqr + q^3.$ 3. Eliminate x by $y = s_2 - x^2.$
5. Eliminate x by $p^2 - q + px = y.$ 7. $2, 4, -6.$
8. $pqr - p^2s - r^2 = 0, \quad x^2 + r/p, \quad x^2 + px + ps/r.$
9. $1, 3, 1 \pm i.$ 12. $z^5 + 2z^4 + 5z^3 + 3z^2 - 2z - 9 = 0.$
13. $z^3 + 15z^2 + 52z - 36 = 0.$ 89. See Ex. 17, p. 78.

INDEX

(The numbers refer to pages.)